



# **STIC Search Report**

## **Biotech-Chem Library**

**STIC Database Tracking Number: 143685**

**TO: Ruixiang Li**  
**Location: rem/4d75/4c70**  
**Art Unit: 1646**  
**Monday, February 14, 2005**

**Case Serial Number: 10/689832**

**From: Edward Hart**  
**Location: Biotech-Chem Library**  
**REM-1A55**  
**Phone: 571-272-2512**

**edward.hart@uspto.gov**

### **Search Notes**

Examiner Li,

Here are the results of the search you requested.

Please feel free to contact me if you have any questions.

Edward Hart



STIC-Biotech/ChemLib

143 685

mg

From: Li, Ruixiang  
Sent: Sunday, January 30, 2005 2:05 PM  
To: STIC-Biotech/ChemLib  
Subject: Sequence search of Application No.10/689,832

RECEIVED  
JAN 31 2005  
(STIC)

Please do a standard search on:

- (i). SEQ ID NOS: 19 and 20 against commercial nucleic acid databases;
- (ii). SEQ ID NO: 20 against commercial amino acid databases.

Thank you very much!

Ruixiang Li  
GAU 1646  
REM 4D75  
Mail Box 4C70  
(571) 272-0875

\*\*\*\*\*

STAFF USE ONLY

Searcher: \_\_\_\_\_  
Searcher Phone: 2- \_\_\_\_\_  
Date Searcher Picked up: 2/8/05  
Date Completed: 2/14/05  
Searcher Prep/Rev. Time: \_\_\_\_\_  
Online Time: \_\_\_\_\_

\*\*\*\*\*

Type of Search 1  
NA Sequence: # \_\_\_\_\_  
AA Sequence: # 1+1 - reverse to NA  
Structure: # \_\_\_\_\_  
Bibliographic: \_\_\_\_\_  
Litigation: \_\_\_\_\_  
Patent Family: \_\_\_\_\_  
Other: \_\_\_\_\_

\*\*\*\*\*

Vendors and cost where applicable  
STN: \_\_\_\_\_  
DIALOG: \_\_\_\_\_  
QUESTEL/ORBIT: \_\_\_\_\_  
LEXIS/NEXIS: \_\_\_\_\_  
SEQUENCE SYSTEM: QSP  
WWW/Internet: \_\_\_\_\_  
Other(Specify): \_\_\_\_\_





GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 10:30:01 ; Search time 3513 Seconds  
(without alignments)  
12124.656 Million cell updates/sec

Title: US-10-689-832-19  
Perfect score: 1119  
Sequence: 1 atggagcacacgacgccca.....cttgtgtcatcatctgtga 1119

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : EST.\*

- 1: gb\_est1.\*
- 2: gb\_est2.\*
- 3: gb\_hic.\*
- 4: gb\_est3.\*
- 5: gb\_est4.\*
- 6: gb\_est5.\*
- 7: gb\_est6.\*
- 8: gb\_gss1.\*
- 9: gb\_gss2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
C 1	457.8	40.9	803	9	CC905917 t025011ba
C 2	250	22.3	362	8	AZ513631 1M0359L07
C 3	229.4	20.5	448	8	AQ225693 HS 2009.B
C 4	218.8	19.6	813	8	BZ159352 CH230-259
C 5	209.8	18.7	806	9	CNS02284 AL178069 Tetradon
C 6	200	17.9	588	8	AQ344035 RPC111-12
C 7	197	17.6	603	8	AQ351433 RPC111-11
C 8	183	16.4	824	9	CL137737 ISB1-110A
C 9	182.8	16.3	947	9	CL137737 ISB1-110A
C 10	173	15.5	708	5	BU610327 UT-M-DJ2
C 11	169.4	15.1	1289	9	CL649540 CH213-237
C 12	150	13.4	867	9	CL137848 ISB1-110C
C 13	149	13.3	898	9	CNS02ENP Tetradon
C 14	111.8	10.0	1107	9	CNS04L7F Tetradon
C 15	111	9.9	515	1	AL921815 AL921815
C 16	87.6	7.8	816	8	AZ535744 ENTQ25TR
C 17	81.6	7.3	891	8	AZ683582 ENTK47TR
C 18	81.6	7.3	906	8	BH153606 ENTW583TF
C 19	77.8	7.0	843	8	AZ551618 ENTW54TR
C 20	77.4	6.9	908	8	AZ548467 ENTPK30TR
C 21	75	6.7	890	8	BH146886 ENTPK48TF
C 22	71.8	6.4	900	8	AZ549980 ENTND94TF
C 23	71	6.3	976	8	BH149983 ENTQD93TF
C 24	70.2	6.3	358	2	BB870915 BB870915

25	69.2	6.2	877	8	AZ531291	AZ531291	ENTBQ34TR
C 26	69.2	6.2	912	8	AZ551092	AZ551092	ENTFJ22TF
C 27	68	6.1	849	8	AZ546009	ENTFW53TF	
C 28	67.8	6.1	488	9	FR0012324	AL003576	F.rubripe
C 29	67.6	6.0	888	8	AZ528430	ENTCO24TR	
C 30	67	6.0	905	8	AZ550256	ENTVE58TR	
C 31	66.2	5.9	721	5	BW141179	BW141179	BW141179
C 32	66	5.9	931	8	BH160272	ENTQV49TR	
C 33	63.6	5.7	890	8	AZ530768	AZ530768	ENTBH54TF
C 34	61.8	5.5	774	9	CNS02APC	AL188841	Tetradon
C 35	61.8	5.5	881	7	CNS97894	CNS97894	TTE000070
C 36	61.4	5.5	607	4	BJ348812	BJ348812	
C 37	61.2	5.5	692	7	CR437869	CR437869	
C 38	61.2	5.5	726	7	CR438428	CR438428	
C 39	61.2	5.5	735	7	CR441503	CR441503	
C 40	61.2	5.5	768	7	CR445667	CR445667	
C 41	60.8	5.4	880	8	AZ529191	AZ529191	ENTBV68TR
C 42	60.8	5.4	954	7	CO250025	CO250025	AGENCOURT
C 43	60.6	5.4	712	7	CR438462	CR438462	
C 44	60	5.4	787	5	BW391073	BW391073	
C 45	60	5.4	1023	9	CNS05805	AL326174	Tetradon

ALIGNMENTS

RESULT 1  
CC905917/c  
LOCUS CC905917 803 bp DNA linear GSS 08-AUG-2003  
DEFINITION t025011ba.x1 TAMBT Bos taurus genomic clone t025011ba, genomic survey sequence.  
ACCESSION CC905917  
VERSION CC905917.1 GI:33524850  
KEYWORDS GSS.  
SOURCE Bos taurus (cow)  
ORGANISM Bos taurus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactylia; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
REFERENCE 1 (bases 1 to 803)  
AUTHORS Lin, S., Najar, F. Z., Adelson, D., Gill, C. A. and Roe, B. A.  
TITLE Bovine BAC End Sequences from Library TAMBT  
JOURNAL Unpublished (2003)  
COMMENT Contact: Bruce A. Roe  
Advanced Center for Genome Technology  
University of Oklahoma Department of Chemistry and Biochemistry  
620 Parrington Oval, Room 208, Norman, OK 73019, USA  
Tel: 405 325 4912  
Fax: 405 325 7762  
Email: broe@ou.edu  
Class: BAC ends  
High quality sequence start: 39  
High quality sequence stop: 551.  
Location/Qualifiers  
source  
1. .803  
/organism="Bos taurus"  
/mol\_type="genomic DNA"  
/strain="Angus bull T A M U Shoshone Y6 11519666"  
/db\_xref="taxon:9913"  
/clone="t025011ba"  
/sex="Male"  
/cell\_type="Blood"  
/clone\_lib="TAMBT"  
/note="Vector: pBelobAC11; Site 1: HindIII; Site 2: HindIII; TAMBT Bovine BAC library (Male) produced by Texas A&M University, Department of Animal Science."

Query Match 40.9%; Score 457.8; DB 9; Length 803;  
Best Local Similarity 88.9%; Pred. No. 2.2e-111;  
Matches 506; Conservative 0; Mismatches 62; Indels 1; Gaps 1;

QY 483 CCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTATCAGCAC-CTCTGTGCATC 541

```

766 CCCTAACAGCTGGGGGCCCAACCCCTGGACTTGAGATTACATCAGCAGGTCATGCATC 707
QY 542 AGCTCCTCATCTGGATTCACACTGCTTCACCGTCTACTGTCGCCCTGCTCCATCTTCTTCA 601
Db 706 ATGTCTCTTATCTGAATTCAGTCTTCCCTGTGTACTGGGTACCCCTGTTCCATATTTCTTCA 647
QY 602 TCTTGAACCTCATATCTGTGTACAGCTCAGGAGAGAGCAATTTTGTCTCCGTGGCT 661
Db 646 TCTGAACTCTCATCATCTGTGTACAGCTCAGGAGAGAGCAATTTTGTCTCCGTGGCT 587
QY 662 ACTCCACGGGGAAGACCCCGCATCTTGTTCACCATTACTTCCATCTTTTGGCCACACTTT 721
Db 586 ACTCCACGGGGAAGACCCCGCATCTTGTTCACCATTACTTCCATCTTTTGGCCACCTCT 527
QY 722 GGGCCCCCGCATCATCATGATCTTTTACCACTCTATGTGGGGCGCCATCCAGAACCGCT 781
Db 526 GGGCCCCCGCATCATCATGATCTTCTACCACTTTTACCGAGCGCCATCCAGAACCGCT 467
QY 782 GSCTGCTACACATCATGTCCGACATTCGCAACATGTCAGCCCTTCTGAACACAGCCATCA 841
Db 466 GSGCTGGTGCATCATGTCCGATGTTCGCAACATGTCAGCCCTTCTGAACACAGCCATCA 407
QY 842 ACTTCTTCTCTACTCTCTTCTCATCAGCAAGCGGTTCCGCAACCATGTCAGCCCGCCAGCTCA 901
Db 406 ACTTCTTCTCTACTCTCTTCTCATCAGCAAGCGGTTCCGTCGATGCGCGCTACTCTCA 347
QY 902 AGGCTTTCTTCAAGTGCAGAGCAAGCACTGTACAGTCTTACACCAATCATAACTTTTCA 961
Db 346 AGGCTTTCTTCAAGTGCAGAGCAAGCACTGTACAGTCTTACACCAATCATAACTTTTCA 287
QY 962 TTAACAGTAGCCCTGGATCTCTCGCGCAACTCACATGTCAGATGCTGGGTACC 1021
Db 286 TTAACAGTAGCCCTGGATCTCTCACGGCAACTCACATGTCAGATGCTGGGTACC 227
QY 1022 AGTATGACAAAAATGGAAAACTATAAAA 1050
Db 226 AGTATGACAAAAATGGAAAACTATAAAA 198

RESULT 2
AZS13631/c
LOCUS
DEFINITION
1M0359107R Mouse 10kb plasmid UUGCLM library Mus musculus genomic
Clone UUGCLM0359107 R, genomic survey sequence.
ACCESSION
AZS13631
VERSION
GSS.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T.,
Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von
Niederhausern,A. and Wright,D.,Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0359 row: L column: 07
Seq primer: CACACGAGAAACAGCTATGACC
Class: plasmid ends
High quality sequence stop: 362.

FEATURES
source
Location/Qualifiers
1..362
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0359107"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/clone_lib="Mouse 10kb plasmid UUGCLM library"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RI. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."
```

## ORIGIN

```

Query Match      22.3%; Score 250; DB 8; Length 362;
Best Local Similarity 91.1%; Pred. No. 9.4e-56;
Matches 265; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 760 GGGGCGCCCATCCAGAACCGCTGGTGTACATCATCTCCGACATTCGCCAACATGTCTA 819
Db 361 GGAGCAGCCATCCAGAACCCCTGGCTGGTGCACATCATGTGGATGTTGCCAACATGTCTA 302
QY 820 GCCCTTCTGAAACAGACGCAATCACTTCTTCTCTACTCTCTTCATCAGCAAGGGTTCGC 879
Db 301 GCCCTTNTGAAACAGACGCAATCACTTCTTCTCTACTCTTTCATCAGCAAGCGTTCGGT 242
QY 880 ACCATGGCAGCGCCAGCTCAAGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTC 939
Db 241 ACCATGGCAGCTGCCCACTCAAGCGCTTGTTCAGTGTCAAGACGACCTGTACAGTTC 182
QY 940 TACACCAATCATAACTTTTCCATAACAGTAGCCCCCTGGATCTCGCCGGCAAACTCACAC 999
Db 181 TATACCAACCATAACTTTTCCATAACAGTAGTCCCTGGATCTCACCAGCAAACTCACAC 122
QY 1000 TGCATCAAGATGCTGGTGTACAGTATGACAAAATGGAACCTATATAAA 1050
Db 121 TGCATCAAGATGCTGGTGTACAGTATGACAAAATGGAACCTATATAAA 71
```

## RESULT 3

AQ225693/c

LOCUS

DEFINITION

HS 2009 B2 B08 T7 CIT Approved Human Genomic Sperm Library D Homo

sapiens genomic clone Plate=2009 Col=16 Row=D, genomic survey

sequence.

ACCESSION

AQ225693

VERSION

AQ225693.1

GI:3650922

KEYWORDS

GSS.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 448)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
MEDLINE 99380589  
PUBMED 10449764  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 2009 row: D column: 16  
Class: BAC ends  
High quality sequence stop: 448.

FEATURES  
source

1. .448  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate:2009 Col=16 Row=D"  
/sex="male"  
/clone\_lib="citr Approved Human Genomic Sperm Library D"  
/notes="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

## ORIGIN

Query Match 20.5%; Score 229.4; DB 8; Length 448;  
Best Local Similarity 93.4%; Pred. No. 3.4e-50;  
Matches 239; Conservative 0; Mismatches 17; Indels 0; Gaps 0;  
QY 795 CATGTCGGACATGCGCAACATGCTAGCCCTTCGAAACAGAGCCATCAACTTCTTCCTCTA 854  
DB 445 CAGGTGACGACATAGCACAGTGTAGCCCTTGTGACACAGCATCACTTCTTCCTATA 386  
QY 855 CTGCTTCATCAGCAAGGGTTCGACACCATGAGCGCCGACGCTCAAGGCTTCTTCAA 914  
DB 385 CTGCGTCATCAGCAAGGGTTCGACACCATGAGCGCCGACGCTCAAGGCTTCTTCAA 326  
QY 915 GTGCCAGAGCACCTGTACAGTCTTACACCATCATTAATTTTCCATACAGTAGCC 974  
DB 325 GTGCCAGAGCACCTGTACAGTCTTACACCATCATTAATTTTCCATACAGTAGCC 266  
QY 975 CTGGATCTCGCCGGCAAACTCACATGTCATCAAGATGCTGGTGATACAGTATGACAAAA 1034  
DB 265 CTGGATCTCGCCGGCAAACTCACATGTCATCAAGATGCTGGTGATACAGTATGACAAAA 206  
QY 1035 TGGAAAACTTATAAA 1050  
DB 205 TGGAAAACTTATAAA 190

## RESULT 4

BZ169352  
LOCUS CH230-259E22.TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone  
DEFINITION CH230-259E22, genomic survey sequence.

## ACCESSION

BZ169352

## VERSION

BZ169352.1 GI:23810547

## KEYWORDS

## SOURCE

Rattus norvegicus (Norway rat)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
Rattus.

## REFERENCE

## AUTHORS

Zhao, S., Shetty, J., Shatman, S., Tseng, G., Geer, K.,  
Shvartsbeyn, A., Gebregeorgis, E., Overton, L., Russell, D., Chen, D.,  
Riggs, P., de Jong, P. and Fraser, C.M.  
Rat BAC End Sequences from Library CHORI-230 MboI segment  
Unpublished (1999)  
Other\_GSSs: CH230-259E22.TV

Contact: Shaying Zhao  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: szhao@tigr.org  
Clones are derived from the rat BAC library CHORI-230  
(http://www.chori.org/bacpac/rat230.htm). For BAC library  
availability, please contact Pieter de Jong (pdejong@mail.cho.org).  
Clones may be purchased from BACPAC Resources  
(http://www.chori.org/bacpac/or ering\_information.htm). BAC end  
page: http://www.tigr.org/todb/bac\_ends/rat/bac\_end\_intro.html  
Plate: 259 row: E column: 22  
Seq primer: SP6  
Class: BAC ends.

FEATURES  
source

1. .813  
Location/Qualifiers  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/strain="BN/SsNHsd/MCW"  
/db\_xref="taxon:10116"  
/clone="CH230-259E22"  
/sex="female"  
/cell\_type="Brain"  
/clone\_lib="CHORI-230 Segment 2"  
/note="Vector: pTARBAC1.3; Site 1: MboI; Site 2: MboI;  
CHORI-230 Rat (BN/SsNHsd/MCW) BAC library produced by  
Pieter de Jong"

## ORIGIN

Query Match 19.6%; Score 218.8; DB 8; Length 813;  
Best Local Similarity 86.7%; Pred. No. 2.8e-47;  
Matches 241; Conservative 0; Mismatches 37; Indels 0; Gaps 0;  
QY 125 CAGCAAAATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGACAGAAAGTCCT 184  
DB 536 CAGCAAAATCTTGACAGTGATCATCTCTCACTGGTAGCAGACAGAGTCCT 595  
QY 185 CCTACAACTATCTTGGCACTCGCTGCGCGACATCTTGGTCTCTTTTCATAGTGT 244  
DB 596 CCTACAACTATCTTGGCACTCGCTGCGCGACATCTTGGTCTCTTTTCATAGTGT 655  
QY 245 TTGTGGACTCTCTGTGGAAGATTTTCATCTTGAACATGACAGATGCTCAGGTCCCGACA 304  
DB 656 TTGTGGATTTCTTGTAGAACTTCATTTTGACCATGACAGATGCTCCGATCCCTGAC 715  
QY 305 AGATCATAGAAGTGTGGAAATTTTCATCCATCCACACCTCCATATGGATACGTACCGT 364  
DB 716 AGATCATAGAAGTGTAGAGTTCTCTCCATCCACACTTCTATTTGGATTACAGTCCCT 775  
QY 365 TAACCATTTGACAGGTATATCGCTGCTGCGCCCGCTC 402  
DB 776 TAACCATTTGATAGGTATATCGCAGTCTGTCAACCCACTC 813

## RESULT 5

## CNS022E4

## LOCUS

## DEFINITION

CNS022E4  
Tetraodon nigroviridis genome survey sequence T7 end of clone  
227D19 of library G from Tetraodon nigroviridis, genomic survey  
sequence.

## ACCESSION

AL178069

## VERSION

AL178069.1 GI:7816126

## KEYWORDS

## SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae;  
1  
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,

Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
10835645  
2  
Roest Crollius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)  
20359837  
10899143  
3 (bases 1 to 806)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr)  
- Web : www.genoscope.cns.fr  
This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.

## FEATURES

source  
1..806  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="227D19"  
/clone\_lib="G"  
/note="Genoscope sequence ID : C0AG227CB10LP1-end : T7"

## ORIGIN

Query Match 18.7%; Score 209.8; DB 9; Length 806;  
Best Local Similarity 56.3%; Pred. No. 7.4e-45;  
Matches 418; Conservative 7; Mismatches 289; Indels 29; Gaps 1;  
QY 166 GCAAGAAGACAGAAGTCTCTTACAACTATCTTTGGCACTCGCTGCTGCCGACATCTTG 225  
DB 2 GGAAAGACCAAGAGCGCTGTACTACTCTGCGGGGGTGACAGGCTCGGACATCTCTC 61  
QY 226 GTCTCTTTTTCATAGTGTGTCGACTCTCTGTTGGAAGATTTTCATCTTGAACATGCG 285  
DB 62 TCCAGCTCTTCATCATCTCTGCTGGCTTCTCTGTTGGAGCGCGGTTTTCACCGGGAG 121  
QY 286 ATGCCCTCAGTCCCGACAGATCATAGAAGTGTGGATTCTCATCCATCCACACCTCC 345  
DB 122 GTCCCGGCTCTCTTACACTCAGTCAGCGCGCGCGAGTTTGGCCGACACACGCTCC 181  
QY 346 ATATGGATTACTGTACCGTTAACCAATTGACAGGTATATGCTGTCTGCCACCCGCTCAAG 405  
DB 182 ATCTGTCTCACCGTCCCCCTCACCGTGGACCGCTAGTGGCGCTGTGCCACCCCTCTCT 241  
QY 406 TACACACGGTCTCATACCCAGCCCGCACCCGGAAGTCATTTGAATGTTTACATCACC 465  
DB 242 CACAGGCAGATCAGCTTACCCGCGCGGACAGGAGATCATCGCGTGTCTCTGTGTG 301  
QY 466 TGCTTCTCAGCAGATCCCTTATTACTGTGTCGCCCAACATCTGACTGAAGACTACATC 525  
DB 302 TCGTGGCTCGGGCTCGCCCTTCTTCTGTGTCCGATCTGAGGAGGAGGAGGAGGAGG 361  
QY 526 AGCACCCTGTGTCATCAGCTCTCTATCTGATCCACTGCTTCCACCGTCTACCTGGTGCC 585  
DB 362 CCGAMGGCGCTGGACGCGCTCTCATCTGGACCCAGTGAACCATCATCTACTTCTTGCC 421  
QY 586 TCGTCCATCTTCTTCATCTTGAATCATCATCTGTTGTTGTTGTTGTTGTTGTTGTTGTT 645  
DB 422 TGCAGCATCTTCTGCTGCTCAACTCTTTGATCATCCMACGCTGAGGCTGGCGGAGAG 481  
QY 646 TTTCTGCTCTCGTGGCTACTCCAC-----GGGGAGAGA 676

DB CGGACGCGCGGGGGCGGCGGCMCAAGTMCACGCTCCGCGCGCTGGGGAAGA 541  
QY 677 CCACGCGCATCTTGTTCACCATTAACCTTCCATCTTTGGCACACTTTGGGCGCCCGCATCA 736  
DB 542 GCACGGCCATGCTGCTGGCCATCACTCCGCTCTTCTGTGTCTCTGGGACCCAGGAGCG 601  
QY 737 TCATGATCTTTTACCACTCTATGGGGCGCCATCCAGACCGCTGGCTGGTGTACATCA 796  
DB 602 TGGTGGTCACTTACCACCTGTACGTGTAGTGGTTCAACGGGAMTGGCGGCTCCACCTGG 661  
QY 797 TGTCCGACATTTGCCAAACATGCTAGCCCTTCTGAACACAGCCATCAACTTCTTCTCTACT 856  
DB 662 CCTAGGACCTGTCCAAACATGCTGGCCATGCTCAACACGCGGCTAACTTCTTCTGTACT 721  
QY 857 GCTTCATCAGCAAGCGGTTCCGC 879  
DB 722 GCTTCGTCAGCAAGCGGTTCCGC 744

## RESULT 6

AQ344035/c  
LOCUS  
DEFINITION  
genomic survey sequence.  
ACCESSION  
AQ344035  
VERSION  
AQ344035.1 GI:4168931  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE  
1 (bases 1 to 588)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and  
Venter, J.C.  
AUTHORS  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Map Building  
Unpublished (1997)  
JOURNAL  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
COMMENT  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbs@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/tldb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: 17

## FEATURES

source  
Location/Qualifiers  
1..588  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7547470"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-124J23"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_lib="RPCI-11"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

## ORIGIN

Query Match 17.9%; Score 200; DB 8; Length 588;  
Best Local Similarity 100.0%; Pred. No. 2.8e-42;  
Matches 200; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 125 CAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCT 184

```
Db 201 CAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCT 142
Qy 185 CCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 244
Db 141 CCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 82
Qy 245 TTGTGGACTTCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACA 304
Db 81 TTGTGGACTTCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACA 22
Qy 305 AGATCATAGAGTGTGGAA 324
Db 21 AGATCATAGAGTGTGGAA 2

RESULT 7
AQ351433/c 603 bp DNA linear GSS 07-MAY-1999
LOCUS RPCI11-113113.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-113113,
DEFINITION genomic survey sequence.
ACCESSION AQ351433
VERSION AQ351433.1 GI:4178768
KEYWORDS GSS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and
Venter, J.C.
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
JOURNAL Map Building
COMMENT Unpublished (1997)
Other GSSs: RPCI11-113113.TV
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html
Seq primer: SP6
Class: BAC ends.
FEATURES
Location/Qualifiers
1..603
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GBB:7543212"
/db_xref="taxon:9606"
/clone="RPCI-11-113113"
/sex="Male"
/cell_type="Lymphocytes"
/clone_lib="RPCI-11"
/notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
RPCI11 Human Male BAC Library"
ORIGIN
Query Match 17.6%; Score 197; DB 8; Length 603;
Best Local Similarity 100.0%; Pred. No. 1.8e-41;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 125 CAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCT 184
Db 197 CAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCT 138
Qy 185 CCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 244
```

```
Db 137 CCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 78
Qy 245 TTGTGGACTTCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACA 304
Db 77 TTGTGGACTTCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACA 18
Qy 305 AGATCATAGAGTGTGCTG 321
Db 17 AGATCATAGAGTGTGCTG 1

RESULT 8
CLI37737
LOCUS ISBI-110A2.T7.1 ISBI Xenopus tropicalis genomic clone ISBI-110A2,
DEFINITION genomic survey sequence.
ACCESSION CLI37737
VERSION CLI37737.1 GI:40631372
KEYWORDS GSS
SOURCE Xenopus tropicalis (western clawed frog)
ORGANISM Xenopus tropicalis
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus; Silurana.
1 (bases 1 to 824)
Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
Mardis, E. and Wilson, R.
TITLE A physical map of the xenopus tropicalis genome
JOURNAL Unpublished (2003)
COMMENT Contact: Richard K Wilson
Genome Sequencing Center
Washington University School of Medicine
Email: submissions@watson.wustl.edu
Insert Length: 75000 Std Error: 0.00
Seq primer: T7 TAATACGACTCACTATAGG
Class: BAC ends
High quality sequence start: 2
High quality sequence stop: 672.
Location/Qualifiers
1..824
/organism="Xenopus tropicalis"
/mol_type="genomic DNA"
/db_xref="taxon:8364"
/clone="ISBI-110A2"
/clone_lib="ISBI"
/notes="Vector: pBelobAC11; ISB-1 Xenopus tropicalis BAC
Library Segment 1"
ORIGIN
Query Match 16.4%; Score 183; DB 9; Length 824;
Best Local Similarity 59.5%; Pred. No. 1.2e-37;
Matches 309; Conservative 0; Mismatches 210; Indels 0; Gaps 0;
Qy 125 CAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCT 184
Db 167 CAGTGAACATGCTGACTGCTGCTGGCACTCTCAAGCTGGCAAGACAGAGTCTCT 226
Qy 185 CCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 244
Db 227 CCTACAGTTACTGTTGGCTCTTACCATCTCGACATCTTGGCCAGATTTTCATCATTT 286
Qy 245 TTGTGGACTTCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACA 304
Db 287 TTGTGGGCTTTATCTCGCAACAGCAATACTCCACGGCAAGGTGCCAGTACCTTAATCC 346
Qy 305 AGATCATAGAGTGTGGAATTTTCATCTCCATCCACCTCCATATGGATTACTGTACCGT 364
Db 347 ATGTGGTCAGTGTCTCTTGGATTTCTCTCTAAATCATGCATCTTATTTGGGTGACAGTATAC 406
Qy 365 TAAACATTGACAGGTATATCGCTGTGCGACCGCTCAAGTACACACAGGTCTCATACC 424
Db 407 TGACCGTGGACCGTTATGTGGCATTTATGTATCCGCTACATATCGCTCTTTCTCTTACC 466
```



/lab\_host="DH108 (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="NIH\_EMAP\_DJ2"  
 /notes="Organ: brain; Vector: pT7T3-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-M-DJ2 is a subtracted cDNA library constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT7T3-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (d)18 tail. The sequence tag for this library is: GCTACATGAT, subfornical organ and area postrema."

## ORIGIN

Query Match 15.5%; Score 173; DB 5; Length 708;  
 Best Local Similarity 92.4%; Pred. No. 5.9e-35;  
 Matches 182; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 854 ACTGCTTCATCAGCAAGCGGTCCGACCAATGCGCCGCGCTCAAGGCTTTCTTCA 913  
 DB 1 ACTGCTTCATCAGCAAGCGGTCCGACCAATGCGCCGCTCAAGGCTTTCTTCA 60

QY 914 AGTGCAGAGCAACTGTACAGTTCTACACCAATCATAACTTTTCCATAAAGTAGCC 973  
 DB 61 AGTGTGAGAAGCAGCGCTGTACAGTTCTATACCAACCACTTCTTCCATAAAGTAGTC 120

QY 974 CTGTGATCTGCGCGGCAACTCAGCTGCATCAAGATGCTGTGACCAATGACAAA 1033  
 DB 121 CTGTGATCTCACCAGCAAACTCAGCTGCATCAAGATGCTGTGACCAATGACAAA 180

QY 1034 ATGGAACCTATAAAA 1050  
 DB 181 ATGGAAGCCTATAAAA 197

## RESULT 11

LOCUS CL649540 1269 bp DNA linear GSS 06-JUL-2004  
 DEFINITION CH213-237F13.SP6 CH213 Gasterosteus aculeatus genomic clone  
 CL649540  
 CL649540.1 GI:49668964  
 GSS.  
 SOURCE Gasterosteus aculeatus (three spined stickleback)  
 ORGANISM Gasterosteus aculeatus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes; Gasterosteidae; Gasterosteus.  
 1 (bases 1 to 1269)  
 Kingley,D., Grimwood,J., Dickson,M., Schmutz,J. and Myers,R.M.  
 Expressed sequence tags from Gasterosteus aculeatus  
 Unpublished (2004)  
 CONTACT: Grimwood, Jane  
 Stanford Human Genome Center  
 Stanford University School of Medicine  
 975 S California Avenue, Palo Alto, CA 94304, USA  
 Tel: 650 320 5917  
 Fax: 650 320 5801  
 Email: jane@shgc.stanford.edu  
 Plate: 237  
 Class: BAC ends  
 High quality sequence start: 19  
 High quality sequence stop: 869.  
 Location/Qualifiers  
 1. .1269  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="genomic DNA"  
 /strain="Salmon River"  
 /db\_xref="taxon:69293"

## FEATURES

## source

1. .1269  
 /organism="Gasterosteus aculeatus"  
 /mol\_type="genomic DNA"  
 /strain="Salmon River"  
 /db\_xref="taxon:69293"

/clone="CH213-237F13"  
 /sex="Mixed"  
 /cell\_type="Blood"  
 /clone\_lib="CH213"

/note="Vector: pTARBAC2.1; Site 1: EcoRI; The sequence of the clone was established as a mapping and sequencing collaboration at the Stanford Genome Evolution Center, funded by the NIH Centers of Excellence in Genomic Science (CEGS) initiative (<http://cegs.stanford.edu>). The clone was isolated from the BAC library CHORI-213 built by Pieter deJong in collaboration with the Stanford Genome Evolution Center (<http://www.chori.org/bacpac/>). Clones may be purchased from BACPAC Resources ([http://www.chori.org/bacpac/ordering\\_information.htm](http://www.chori.org/bacpac/ordering_information.htm))."

## ORIGIN

Query Match 15.1%; Score 169.4; DB 9; Length 1269;  
 Best Local Similarity 77.2%; Pred. No. 5.9e-34;  
 Matches 206; Conservative 0; Mismatches 61; Indels 0; Gaps 0;

QY 777 CGCGTGGCTGTACATCATGTCCGACATTCGCAATTCGCAATCTAGCCCTTCTGAACACAGC 836  
 DB 42 CGGGCGACTGCTCCAGTGCTCACCAGCTGCGCATATATGCTCGCGTTGCTCAACACCGG 101

QY 837 CATCAACTTCTTCTTACTGCTTATCAGAGCGGTTCCGACCATGCGAGCCGCCAC 896  
 DB 102 GGTCAACTTCTTCTTACTGCTTATCAGCAAGCGTTTCCGCGGCATGCGGCCAACGT 161

QY 897 GCTCAAGCTTCTTCAAGTGCAGCAAGCAACTGTGTACAGTTCTACACCAATCATACTT 956  
 DB 162 CTGCGGGCGCTTGGTCACTGCGGGAAGCAGCGCGCGTCTACGCCAGCCACAACTT 221

QY 957 TTCCCAATAACAAGTAGCCCTGATCTCGCGGCAAACTCACACTGTCATCAAGATGCTGGT 1016  
 DB 222 TTCCATCAACAGCAGCGCGTGTATCTACCGGCCAACTCCCACTGCTTAAGATGCTGGT 281

QY 1017 GTACCAGTATGACAAAATGGAAACC 1043  
 DB 282 GTACCAGTATGACAAAATGGAAACC 308

## RESULT 12

LOCUS CL137848 867 bp DNA linear GSS 05-JAN-2004  
 DEFINITION ISB1-110C15 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15, genomic survey sequence.  
 CL137848  
 CL137848.1 GI:40631483  
 GSS.  
 SOURCE Xenopus tropicalis (western clawed frog)  
 ORGANISM Xenopus tropicalis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus; Silurana.  
 1 (bases 1 to 867)  
 Kremitzki,C., Carter,J., McPherson,J., Warren,M., Graves,T., Mardis,E. and Wilson,R.  
 A physical map of the xenopus tropicalis genome  
 Unpublished (2003)  
 CONTACT: Richard K Wilson  
 Genome Sequencing Center  
 Washington University School of Medicine  
 Email: [submissions@watson.wustl.edu](mailto:submissions@watson.wustl.edu)  
 Insert Length: 75000 Std Error: 0.00  
 Seq primer: T7 TAATACGACTCACTATAGG  
 Class: BAC ends  
 High quality sequence start: 11  
 High quality sequence stop: 673.  
 Location/Qualifiers  
 1. .867  
 /organism="Xenopus tropicalis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:8364"

## FEATURES

## source

1. .867  
 /organism="Xenopus tropicalis"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:8364"

/clone="ISB1-110C15"  
/clone\_lib="ISB1"  
/note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC  
Library Segment 1"  
ORIGIN

Query Match 13.4%; Score 150; DB 9; Length 867;  
Best Local Similarity 58.1%; Pred. No. 8.3e-29;  
Matches 264; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 125 CAGCAATATCTTGACAGTATCATCTCCAGCTGTGGCAGAAAGACAGATGCTT 184  
DB 182 CAGTGAACATGCTGACTGTGGCAGCTCTCAAGGCTGCATCGAAGCAAAAGATGCTT 241  
QY 185 CTTACAACTATCTTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATAGTGT 244  
DB 242 CTTACAGTTACTGTGGCTCTTACCATCTCAGACATCTTGAGCCAGATTTTCATCATTT 301  
QY 245 TTGTGGACTTCTGTGGGAAGATTTTCATCTTGAACATGCAGATGCTTCAGGTCCCGGACA 304  
DB 302 TTGTGGCTTTATCTCGCAACAGCAATATCTCCAGCAAGTGCCAGTACCTTAATCC 361  
QY 305 AGATCATAGAAGTGTGGAATTTCTATCTTGAACATGCAGATGCTTCAGGTCCCGGACA 364  
DB 362 ATGTGGTCAAGTGTCTTGAATTTCTCTTAATCATGATCTATTTGGGTGACAGTGATAC 421  
QY 365 TAACCATTCACAGGTATATCGCTGCTGCCACCGCTCAAGTACACACGGTCTCATACC 424  
DB 422 TGACCGTGGACCGTATGTGGCATATGCTATCCGCTACAATATCGCTCTTTCTTTACC 481  
QY 425 CAGCCCGCACCCGGGAAGTCAATGTAAGTGTATCATCACCTGCTTCTGACACGATCC 484  
DB 482 CAGAGCGAACCGTAGAGTATGTTGTCTTCTCTCATCGTTGTGACCGGTGTAC 541  
QY 485 CTTATTAAGTGTGGCCCAACATCTGCACTGAAGATATCATATGACACCTCTGTGCAATCAG 544  
DB 542 CTTCTACTGTGGAGTGTATGTTGGAGGGATCCGCTACCCAGGATGCTTAGACCTTA 601  
QY 545 TCCTCATCTGGATCCACTGCTTCACCGTCTACT 578  
DB 602 TACTCAAGTGGACCCACTGCTTTATATATACTT 635

CNS02ENP 898 bp DNA linear GSS 01-SEP-2000  
Tetraodon nigroviridis genome survey sequence PUC-ori end of clone  
262F08 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
AL193966  
AL193966.1 GI:7832072  
GSS: genome survey sequence.  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontidae; Tetraodontidae; Tetraodon.  
1  
Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C.,  
Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F.,  
Saurin, W. and Weissenbach, J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
10835645  
2  
Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C.,  
Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F.,  
Saurin, W., Bernot, A. and Weissenbach, J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis

Genome Res. 10 (7), 939-949 (2000)  
20359837  
10899143  
3 (bases 1 to 898)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqrefgenoscope.cns.fr)  
- Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/Tetraodon.  
Location/Qualifiers  
1..898  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="262F08"  
/clone\_lib="G"  
/note="Genoscope sequence ID : C0AG262DC04SP1-end :  
PUC-ori"  
ORIGIN

Query Match 13.3%; Score 149; DB 9; Length 898;  
Best Local Similarity 56.9%; Pred. No. 1.6e-28;  
Matches 292; Conservative 0; Mismatches 220; Indels 1; Gaps 1;

QY 121 TTACCAGCAATATCTTGACAGTATCATCTTCCAGCTGTGGTGGCAAGACAGAG 180  
DB 370 TTCCAGTGAACATCTGACGCGGTGGCTTGACCCGCTGGCTCCGCCACCAAG 429  
QY 181 TCCTCTCAACATATCTTGGCACTCGCTGTGCGGACATCTTGGTCTCTTTTCATA 240  
DB 430 GCACGTGATCTACTACCTGTGCGGTGACAGCTCCGACATCTCTCCAGCTCTTCATC 489  
QY 241 GTGTTGTGGACTTCTGTGGAGATTTTCATCTTGAACATGCAGATGCTTCAGTCCCC 300  
DB 490 ATCTTCGTGGGTCTCTGTGGAGACGGCAGTTTTCACCGGAGGTCCCGCGCTCTC 549  
QY 301 GACAAGATCATAGAAGTCTGGAAATTTTCATCCATCCACACCTCCATATGATTA 360  
DB 550 TTACACTCAGTCAGCGCGCGAGTTTGCGCCACCAACGCTCCATCTGGTCCACCGTC 609  
QY 361 CGGTAAACATTCAGAGTATATCGCTGTGCGCACCCGCTCAAGTACACACGCTCA 420  
DB 610 CCCCTCACCGTGGACCGCTACGTGCGCTGTGCCACCCCTCTCCACAGGAGATCAGC 669  
QY 421 TACCCAGCCCGCACCCGGAAAGTCAATGTAAGTGTTCATACATCACCTGCTTCTGACCCAGC 480  
DB 670 TACCGGCGCGGACACAGAGGATCATCGCGGTGTCTGGTGTCTGCGCTGGCGCTGGGC 729  
QY 481 ATCCCTCTATTAAGTGTGGCCCAACATCTGAGTCTGAAGCTATACATCAGACCTCTGTG 540  
DB 730 CTGCGCTCTTCTTGTGGTGTCCGACATGTGGAGGAACAGCCACC -CGCCGACGSGTGA 788  
QY 541 CAGTCTCTCATCTGGATCCATCTGCTTACCTTACCTGTGGTGGCTCTCCATCTTCTTC 600  
DB 789 CGCGTCTCATCTGGACCCACGTGACCATCATCTACTTCTGCGCTGAGCATCTTCTG 848  
QY 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGG 633  
DB 849 GTCTCAACTCTTTGATCATCTCCACGCTGAGG 881

RESULT 14  
CNS04L7P  
LOCUS  
DEFINITION  
Tetraodon nigroviridis genome survey sequence I7 end of clone  
118A07 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
AL295764  
AL295764.1 GI:8034344  
CNS04L7P 1107 bp DNA linear GSS 01-SEP-2000  
Tetraodon nigroviridis genome survey sequence I7 end of clone  
118A07 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
AL295764  
AL295764.1 GI:8034344



KEYWORDS  
SOURCE GSS; genome survey sequence.  
ORGANISM Tetraodon nigroviridis

REFERENCE  
AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Pizames, C., Wincker, P., Brottier, P., Quetier, P., Saurin, W., and Weissenbach, J.  
TITLE Estimate of human gene number provided by genome-wide analysis using Tetraodon nigroviridis DNA sequence  
JOURNAL Nat. Genet. 25 (2), 235-238 (2000)  
MEDLINE 20296633  
PUBMED 10835645

REFERENCE  
AUTHORS Roest Crolius, H., Jaillon, O., Dasilva, C., Orzouf-Costaz, C., Pizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, P., Saurin, W., Bernot, A., and Weissenbach, J.  
TITLE Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis  
JOURNAL Genome Res. 10 (7), 939-949 (2000)  
MEDLINE 20359837  
PUBMED 10899143

REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (12-APR-2000) Genoscope - Centre National de Sequencage : BP 191 9106 EVRY cedex - FRANCE (E-mail : seqrf@genoscope.cns.fr)  
JOURNAL - Web : www.genoscope.cns.fr

COMMENT This sequence is a single read and was generated as part of a large scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/Tetraodon>.

FEATURES  
source  
1. .1107  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="118A07"  
/clone\_lib="G"  
/notes="Genoscope sequence ID : C0BG118AA04LP1-end : T7"

ORIGIN  
Query Match 10.0%; Score 111.8; DB 9; Length 1107;  
Best Local Similarity 70.6%; Pred. No. 1.6e-18;  
Matches 149; Conservative 0; Mismatches 62; Indels 0; Gaps 0;

QY 669 GGGGAAGACACCGCCATCTGTTCACCATTAACCTCCATCTTTGCCACACATTTGGGCCCC 728  
|||||  
Db 70 GGGGAAGACACGGCCATGCTGTGGCCATCACTCCGCTTCTCTGTGCTCTGGGCACC 129  
|||||

QY 729 CGCATCATCATGATTTTACACCTCTATGGGGGGCCCATCCAGAACCGGTGGCTGGT 788  
Db 130 CAGGACGGTGGTGTATCTACCATGTACGTGCTGCGTTACCGGACCTGGCGCT 189  
|||||

QY 789 ACACATCATGTCGCATTCGCAACATGCTAGCCCTTCTGAACACAGCCATCACTTCTT 848  
Db 190 CCACCTGGGCTAGCACTGTCCAAACATGCTGGCCATGCTCAACACGGCGTCACTTCTT 249  
|||||

QY 849 CCTTACTGCTTCATCAGCAAGCGGTTCGGC 879  
Db 250 CCTGTACTGCTTCGTGACGACGCGTTCGGC 280  
|||||

RESULT 15  
LOCUS AL921815  
DEFINITION AL921815 PUR-Z1+22 Danio rerio cdna clone 101-P03-2, mRNA sequence.  
ACCESSION AL921815  
VERSION AL921815.1 GI:23188395  
KEYWORDS EST.

SOURCE  
ORGANISM Danio rerio (zebrafish)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Osteichthyes; Cypriniformes; Cyprinidae; Danio.  
REFERENCE 1 (bases 1 to 515)  
AUTHORS Lo, J., Lee, S., Xu, M., Liu, P., Ruan, H., Sun, A., He, Y., Ma, W., Wang, W., Wen, Z., and Peng, J.  
TITLE 15000 unique zebrafish EST clusters and their future use in microarray for profiling gene expression patterns during embryogenesis  
JOURNAL Genome Res. 13 (3), 455-466 (2003)  
MEDLINE 22505427  
PUBMED 12618376  
COMMENT Contact: Peng J  
Lab of Functional Genomics  
Institute of Molecular and Cell Biology  
30 Medical Drive, Singapore, 117609, Singapore  
Email: pengj@imcb.a-star.edu.sg  
Clone requests: info@openbiosystems.com  
Open Biosystems,  
6705 Odyssey Drive, Huntsville, AL 35806.  
FEATURES  
source  
1. .515  
/organism="Danio rerio"  
/mol\_type="mRNA"  
/strain="local wildtype"  
/db\_xref="taxon:7955"  
/clone="101-P03-2"  
/tissue\_type="whole embryo or fish"  
/dev\_stage="mixed stages"  
/clone\_lib="PUR-Z1+22"

ORIGIN  
Query Match 9.9%; Score 111; DB 1; Length 515;  
Best Local Similarity 63.5%; Pred. No. 2.1e-18;  
Matches 190; Conservative 0; Mismatches 100; Indels 9; Gaps 1;

QY 640 AGCAATTTTCGTCTCGTGGCTACTCCAGGGGAGACACCGCCATCTTTGCCACATT 699  
|||||  
Db 214 AGCTGCTCCGGCTCGAGGATCTCGACCGAAGACACCGGTATCTTTAGCCATC 273  
|||||

QY 700 ACCTCCATCTTTGCCACATTTTGGGCCCCCGCATCATCATGATTTTACCACCTCTAT 759  
|||||  
Db 274 ACCTCAGTGTTCGCGCTTTTATGGGCGCTCGTACACTTATGATTTCTTATCATCTTTAC 333  
|||||

QY 760 GGGGGCCCATCCAGAACCGCTGGCTGG-----TACATCATGTCCGACATGCC 810  
|||||  
Db 334 ACGGTGCAACACGCGATGCGGGTCCGGCTAGACTCCTGCTATTTGGTAACGGATGTAGCG 393  
|||||

QY 811 AACATGTAGCCCTTCTGAACACAGCCATCAACTTCTCTCTACTGCTTTCATCAGCAAG 870  
|||||  
Db 394 AACATGTGCTTCTACTCAACACCGGGTCAACTTCTCTCTACTGCTTTCATCAGCAAG 453  
|||||

QY 871 CGGTTCCGACCATCCAGCCCGCCAGCTCAAGGCTTTCTTCAAGTGCAGAGCAAC 929  
Db 454 CGCTTTCGAGGATGCGCGGACGGTCTCAAGGCTTTTTCGGCTGCAGAAACAGCC 512  
|||||

Search completed: February 12, 2005, 13:28:37  
Job time : 3517 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 08:28:56 ; Search time 627 Seconds  
(without alignments)  
10564.897 Million cell updates/sec

Title: US-10-689-832-19  
Perfect score: 1119  
Sequence: 1 atggagcacgcacgcacca.....cttgtctatcctcgtgta 1119

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_16Dec04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002as:\*  
7: Geneseqn2002bs:\*  
8: Geneseqn2003as:\*  
9: Geneseqn2003bs:\*  
10: Geneseqn2003cs:\*  
11: Geneseqn2003ds:\*  
12: Geneseqn2004as:\*  
13: Geneseqn2004bs:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1119	100.0	1119	4 AAS15730	Aas15730 DNA encod
2	1119	100.0	1119	10 ADJ87766	Adj87766 G-coupled
3	1119	100.0	1119	12 ADI79322	Adi79322 NOV10a co
4	1119	100.0	1119	12 ADO56001	Ado56001 DNA encod
5	1119	100.0	1130	6 AAD29672	Aad29672 Human G-p
6	1119	100.0	2198	9 ACF05275	Acf05275 Human G-p
7	1091	97.5	1110	9 ACF05276	Acf05276 Human G-p
8	1091	97.5	2189	10 ADD18120	Add18120 Human G-p
9	1050	93.8	1059	6 ABN84269	Abn84269 Human che
10	1050	93.8	1062	6 ABK15562	Abk15562 cDNA enco
11	1050	93.8	1062	10 ACC71785	Acc71785 Human G p
12	1050	93.8	1202	8 ABV73364	Abv73364 Human TGR
13	1048.4	93.7	1062	6 ABK15563	Abk15563 cDNA enco
14	1048.4	93.7	1062	6 ABT04873	Abt04873 Human G p
15	1048.4	93.7	1062	10 ACC44115	Acc44115 Human AXO
16	1048.4	93.7	1343	4 AAS15731	Aas15731 DNA encod
17	1048.4	93.7	1343	10 ADJ87768	Adj87768 G-coupled
18	1048.4	93.7	1343	12 ADI79324	Adi79324 NOV11 cod
19	1048.4	93.7	1343	12 ADO56003	Ado56003 DNA encod
20	1048.4	93.7	1776	10 ADF70587	Adf70587 Orphan re

21	1048.4	93.7	2117	13 ADR16433	Adr16433 Human KOR
22	1048.4	93.7	2273	12 ADO28955	Ado28955 Human nov
23	1040.4	93.0	1526	6 AAD46858	Aad46858 Human 7TM
24	1040.4	93.0	1526	8 ACA60889	Aca60889 Human CDN
25	1040.4	93.0	1526	10 ABS57024	Abes57024 cDNA enco
26	934.8	83.5	1062	6 ABN84273	Abn84273 Human che
27	929.8	83.1	957	12 ADI79347	Adi79347 NOV10b co
28	929.8	83.1	957	12 ADO56026	Ado56026 DNA encod
29	926	82.8	1002	6 AAD27501	Aad27501 Human G-p
30	926	82.8	1002	10 AAD61658	Aad61658 Human GPC
31	926	82.8	1158	6 ABZ42880	Abz42880 Human GPC
32	926	82.8	1167	13 ADO89945	Ado89945 Antagonis
33	924.8	82.6	1000	6 AAD28102	Aad28102 Human thy
34	922.4	82.4	1032	6 ABN84271	Abn84271 Human che
35	922.4	82.4	1070	6 ABN84270	Abn84270 Human che
36	922.4	82.4	1826	6 ABN84272	Abn84272 Human che
37	918.8	82.1	946	10 ADJ87879	Adj87879 G-coupled
38	844.2	75.4	864	4 AAS42811	Aas42811 Human G p
39	832	74.4	1038	8 ABV73373	Abv73373 Mouse TGR
40	832	74.4	1038	12 ADO28957	Ado28957 Mouse nov
41	830	74.2	867	6 AAS98134	Aas98134 Human DNA
42	830	74.2	894	6 AAS98143	Aas98143 Human DNA
43	811.2	72.5	963	10 ADC12695	Adc12695 Human GPC
44	756.4	67.6	795	6 AAS98071	Aas98071 Human DNA
45	756.4	67.6	795	8 ABZ42552	Abz42552 Human G p

ALIGNMENTS

RESULT 1  
AAS15730  
ID AAS15730 standard; cDNA; 1119 BP.

AC AAS15730;

DT 14-FEB-2002 (first entry)

DE DNA encoding chemokine receptor family related protein, NOV10.

XX NOV; cytostatic; psoriasis; nootropic; neuroprotectant;  
XX NOV; cytoprotective; hepatic; antiinflammatory; diabetes; anaemia;  
XX haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;  
XX differentiation; proliferation; haematopoiesis; wound healing;  
XX angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;  
XX haemophilia; allergy; Pendred syndrome; skeletal dysplasia;  
XX ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;  
XX chemokine receptor; chromosome 1; 88.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 1..1119

FT /tag= a

FT sig\_peptide /product= "Chemokine receptor related protein, NOV10"

FT /tag= b

FT mat\_peptide 142..1116

FT /tag= c

FT /note= "Mature chemokine receptor related protein, NOV10"

WO200170978-A2.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US009093.

XX 20-MAR-2000; 2000US-0190768P.

XX 20-MAR-2000; 2000US-0190835P.

XX 22-MAR-2000; 2000US-0190972P.

XX 22-MAR-2000; 2000US-0191199P.

XX 24-MAR-2000; 2000US-0191947P.

XX 28-MAR-2000; 2000US-0192657P.

PR 28-MAR-2000; 2000US-0192664P.  
PR 28-MAR-2000; 2000US-0192665P.  
PR 28-MAR-2000; 2000US-0192984P.  
PR 29-MAR-2000; 2000US-0192836P.  
PR 31-MAR-2000; 2000US-0193843P.  
XX (CURA-) CURAGEN CORP.  
XX  
PI Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;  
FI Vernhet CAM;  
XX  
DR WPI; 2001-639127/73.  
DR P-PSDB; AAU10067.  
XX  
XX Polypeptides and nucleic acids related to chloride channel, insulin-like  
PT growth factor family of proteins, useful for diagnosing and treating  
PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.  
XX  
XX Claim 9; Page 42; 151pp; English.  
XX  
CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX  
CC polypeptides are useful for treating pathology associated with NOVX  
CC polypeptide, determining the presence of or predisposition to a disease  
CC associated with altered levels of NOVX, identifying agents binding to  
CC NOVX and treatment of disorders associated with altered expression of  
CC members of chloride channel-associated proteins e.g. cystic fibrosis and  
CC congenital myotonia. NOVX proteins are useful in treatment of disorders  
CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,  
CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells  
CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,  
CC skeletal dysplasias, disorders characterised by altered cell shape,  
CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial  
CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the  
CC treatment of disorders of vascular smooth muscle cell differentiation,  
CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are  
CC useful to screen for molecules which inhibit or enhance NOVX activity or  
CC function and are useful as targets for the identifying small molecules,  
CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell  
CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV  
CC sequences are also useful for: identifying a cell or tissue type in a  
CC biological sample; amplifying DNA sequences from very small biological  
CC samples e.g. hair or skin or body fluids and as primers and probes to  
CC identify and/or clone NOVX homologues. NOVX proteins are useful  
CC immunogens to generate antibodies to monitor protein levels and modulate  
CC NOVX activity. Cells comprising the nucleic acids are useful for  
CC producing transgenic animals, for studying the function and/or activity  
CC of NOVX protein and identifying and/or evaluating modulators of NOVX  
CC protein activity. This sequence encodes NOV10 (located on chromosome 1)  
CC related to the chemokine receptor family of proteins, one of 12 NOV  
CC polypeptides described in the method of the invention  
XX  
XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;  
SQ  
Query Match 100.0%; Score 1119; DB 4; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 1.3e-306;  
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGACACACGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGTGTGTCCTCCCGGC 60  
DB 1 ATGAGACACACGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGTGTGTCCTCCCGGC 60  
QY 61 TCGGCTCGCGCTTGGTTCGTGCGGTCTACTACAGCCTCTTGTGTGTCCTCGGT 120  
DB 61 TCGGCTCGCGCTTGGTTCGTGCGGTCTACTACAGCCTCTTGTGTGTCCTCGGT 120  
QY 121 TTACACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAGAGACAGAAG 180  
DB 121 TTACACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAGAGACAGAAG 180  
QY 181 TCTCTCTACAATATCTTGTGGACATCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240  
DB 181 TCTCTCTACAATATCTTGTGGACATCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240

QY 241 GTGTTTGTGACTTCTCTGTGAAAGATTTCATCTTGAACATGACAGATGCCTCAGGTCCCC 300  
DB GTGTTTGTGACTTCTCTGTGAAAGATTTCATCTTGAACATGACAGATGCCTCAGGTCCCC 300  
QY 301 GACAAGATCATAGAAGTGTCTGGAATTCTCATCCATCCACACCTCCATATGATGATCTGTA 360  
DB GACAAGATCATAGAAGTGTCTGGAATTCTCATCCATCCACACCTCCATATGATGATCTGTA 360  
QY 361 CCGTTAACCATTTGACAGGTATATCGCTGTCTGCGACCCCGCTCAAGTACACACGGTCTCA 420  
DB CCGTTAACCATTTGACAGGTATATCGCTGTCTGCGACCCCGCTCAAGTACACACGGTCTCA 420  
QY 421 TACCCAGCCCGCACCCCGAAAGTCAATTGTAAGTGTTTACATCACCTGCTTCTGACACG 480  
DB TACCCAGCCCGCACCCCGAAAGTCAATTGTAAGTGTTTACATCACCTGCTTCTGACACG 480  
QY 481 ATCCCTTATTAATCTGGTGGCCCAACATCTGGAATGAAGACTACATCAGCACCTCTGTGAT 540  
DB ATCCCTTATTAATCTGGTGGCCCAACATCTGGAATGAAGACTACATCAGCACCTCTGTGAT 540  
QY 541 CAGTCTCTCATCTGGATCCACTGCTTACCTGCTTACCTGCTGCTGCTGCTGCTTCTTTC 600  
DB CAGTCTCTCATCTGGATCCACTGCTTACCTGCTTACCTGCTGCTGCTGCTTCTTTC 600  
QY 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGGAGGAGCAATTTTGTCTCTCGTGGC 660  
DB ATCTTGAACCTCAATCATTTGTGTACAGCTCAGGAGGAGCAATTTTGTCTCTCGTGGC 660  
QY 661 TACTCCACGGGGAAGACACCGCCATCTTGTTCACCAATTAATCTGATCTTTGGCCACATT 720  
DB TACTCCACGGGGAAGACACCGCCATCTTGTTCACCAATTAATCTGATCTTTGGCCACATT 720  
QY 721 TGGGCCCCCGCATCATCATGATCTTTTACCACTCTATGCGGCGCCCATCCAGAACCGC 780  
DB TGGGCCCCCGCATCATCATGATCTTTTACCACTCTATGCGGCGCCCATCCAGAACCGC 780  
QY 781 TGGCTGTGTACATCATCTGCGACATTCGCCAACATGCTAGCCCTTCTGAAACACAGCCATC 840  
DB TGGCTGTGTACATCATCTGCGACATTCGCCAACATGCTAGCCCTTCTGAAACACAGCCATC 840  
QY 841 AACTTCTTCTCTACTGTCTTTCATCAGCAAGCGGTTCCGACCATGCGGACGCGCCACGCTC 900  
DB AACTTCTTCTCTACTGTCTTTCATCAGCAAGCGGTTCCGACCATGCGGACGCGCCACGCTC 900  
QY 901 AAGGCTTTCTCAAGTGCAGAACCTGTACAGTTCTACAGTCTTACCAATTAATCTTTTCC 960  
DB AAGGCTTTCTCAAGTGCAGAACCTGTACAGTTCTACAGTCTTACCAATTAATCTTTTCC 960  
QY 961 ATAAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTAC 1020  
DB ATAAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTGTAC 1020  
QY 1021 CAGTATGACAAAAATGAAAAACCTATAAAAAAGTGTATATGACAGAAAAAGTCTTACCAG 1080  
DB CAGTATGACAAAAATGAAAAACCTATAAAAAAGTGTATATGACAGAAAAAGTCTTACCAG 1080  
QY 1081 TTTGAAGATGCCATTTGGAGCTTGTGTATCTTCTGTGA 1119  
DB TTTGAAGATGCCATTTGGAGCTTGTGTATCTTCTGTGA 1119  
RESULT 2  
ADJ87766  
ID ADJ87766 standard; DNA; 1119 BP.  
XX AC ADJ87766;  
XX AC ADJ87766;  
DT 06-MAY-2004 (first entry)  
DE G-coupled protein receptor-related protein coding sequence #56.  
XX novel protein; G-coupled protein receptor-related protein;  
KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;

KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;  
 KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;  
 KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;  
 KW viral infections; bacterial infection; parasitic infection;  
 KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;  
 KW Alzheimer's disease; tuberous sclerosis; hypercalcaemia; cerebral palsy;  
 KW gene; ds.

XX Unidentified.

XX WO2002102321-A2.

XX 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019522.

XX 18-JUN-2001; 2001US-0298994P.

XX 18-JUN-2001; 2001US-0299134P.

XX 04-OCT-2001; 2001US-00972446.

XX 06-JUN-2002; 2002US-00299134.

XX 07-JUN-2002; 2002US-00298994.

XX (CURA-) CURAGEN CORP.

XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;

XX Pena CE, Spaderna SK, Zhong M;

XX WPI; 2003-167441/16.

XX P-PSDB; ADJ87767.

XX New MOLX polypeptides and polynucleotides, useful in gene therapy,

XX particularly for treating or preventing e.g. cardiomyopathy,

XX atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate

XX cancer, stroke or pain.

XX Claim 8; SEQ ID NO 201; 378pp; English.

XX The invention comprises the amino acid and coding sequences of novel G-

XX coupled protein receptor-related (MOL) proteins. The DNA and protein

XX sequences of the invention are useful for treating or preventing a MOL-

XX associated disorder, such as: cardiomyopathy, atherosclerosis, disorders

XX associated with cell signal processing and metabolic pathway modulation,

XX or diabetes. The DNA and protein sequences are also useful for the

XX treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,

XX pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,

XX systemic lupus erythematosus, viral infections, bacterial infections,

XX parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau

XX syndrome, Alzheimer's disease, tuberous sclerosis, hypercalcaemia, or

XX cerebral palsy. The present DNA sequence encodes a MOL protein of the

XX invention.

XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;

XX Query Match 100.0%; Score 1119; DB 10; Length 1119;

XX Best Local Similarity 100.0%; Pred. No. 1.3e-306;

XX Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

QY

DB

241 GTGTTTGGACTTCTCTGTGGAAGATTTCATCTTGAATGAGATGCTCAGGTCCCC 300

241 GTGTTTGGACTTCTCTGTGGAAGATTTCATCTTGAATGAGATGCTCAGGTCCCC 300

301 GACAAAGATCATAGAAGTCTGGAATTTCTCATCCATCCACACTCCATATGGATTACTGTA 360

301 GACAAAGATCATAGAAGTCTGGAATTTCTCATCCATCCACACTCCATATGGATTACTGTA 360

361 CCGTTAAACCAATTGACAGGTATATCGCTGTCTGCCACCCCGCTCAAGTACCACAGGTCTCA 420

361 CCGTTAAACCAATTGACAGGTATATCGCTGTCTGCCACCCCGCTCAAGTACCACAGGTCTCA 420

421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTATACATCACCTGCTCTCTGACCAGC 480

421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTATACATCACCTGCTCTCTGACCAGC 480

481 ATCCCTTATTACTGTGGGCCCAACATCTGAGTCTGAAGACTACATCAGACCACTCTGTGCAT 540

481 ATCCCTTATTACTGTGGGCCCAACATCTGAGTCTGAAGACTACATCAGACCACTCTGTGCAT 540

541 CACGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGTCCCTGCTCCATCTTCTTC 600

541 CACGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGTCCCTGCTCCATCTTCTTC 600

601 ATCTTGAACCTCAATCATCTGTGTACAAAGCTCAGAGAGAGAGCAATTTTCGTCTCCGTGGC 660

601 ATCTTGAACCTCAATCATCTGTGTACAAAGCTCAGAGAGAGAGCAATTTTCGTCTCCGTGGC 660

661 TACTCCAGGGGAAGACCCGCCATCTGTTTACCACTTACCTCCATCTTCCGACACATT 720

661 TACTCCAGGGGAAGACCCGCCATCTGTTTACCACTTACCTCCATCTTCCGACACATT 720

721 TGGGCCCCCGCGCATCATCATGATTCTTTACCACTCTATGGGGCGGCCCATCCAGAACCGC 780

721 TGGGCCCCCGCGCATCATCATGATTCTTTACCACTCTATGGGGCGGCCCATCCAGAACCGC 780

781 TGGCTGGTACACATCATGTCCGACATGTCGCAACATGCTAGCCCTTCTGAAACACAGCATC 840

781 TGGCTGGTACACATCATGTCCGACATGTCGCAACATGCTAGCCCTTCTGAAACACAGCATC 840

841 AACTTCTTCTCTACTGCTTCTATGAGCAAGCGGTTCCGACCATGGCAGCCGCGCATC 900

841 AACTTCTTCTCTACTGCTTCTATGAGCAAGCGGTTCCGACCATGGCAGCCGCGCATC 900

901 AAGGCTTTCTTCAAGTGGCCAGAACCTGTACAGTGTCTACCACTATCAATCAATCTTTC 960

901 AAGGCTTTCTTCAAGTGGCCAGAACCTGTACAGTGTCTACCACTATCAATCAATCTTTC 960

961 ATAAACAAGTAGCCCTGGATCTCGCCGCGCAAACTCACACTGCATCAAGATGCTGGGTAC 1020

961 ATAAACAAGTAGCCCTGGATCTCGCCGCGCAAACTCACACTGCATCAAGATGCTGGGTAC 1020

1021 CAGTATGCAAAAATGGAACCTTATAAAGTCTTAATGACAGCAAAAGCTCTCTACAG 1080

1021 CAGTATGCAAAAATGGAACCTTATAAAGTCTTAATGACAGCAAAAGCTCTCTACAG 1080

1081 TTTGAAGATGCCATTGGAGCTTGTGTCATCATCTCTGTGA 1119

1081 TTTGAAGATGCCATTGGAGCTTGTGTCATCATCTCTGTGA 1119

RESULT 3

ADIT9322

ID ADIT9322 standard; DNA; 1119 BP.

XX ADIT9322;

XX 22-APR-2004 (first entry)

XX NOV10a coding sequence, SEQ ID 19.

XX Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;

XX Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;

KW

KW Ophthalmological; Antipsoriatic; Neuroprotective; Nootropic;  
 KW Antiarteriosclerotic; Hypotensive; Cardiant; Cerebroprotective;  
 KW Gene Therapy; NOVX; human; cancer; myelogenous leukaemia;  
 KW congenital neonatal autoimmune thrombocytopaenia; immunological disorder;  
 KW allergy; infection; asthma; lung disease; reproductive disorder;  
 KW haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;  
 KW diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;  
 KW hypertension; stroke; heart failure; chromosome 1; NOV10a;  
 KW chemokine receptor; gene; db.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1119  
 FT /\*tag= a  
 FT /product= "NOV10a"  
 XX  
 PN WO2004009635-A2.  
 XX  
 XX 29-JAN-2004.  
 XX  
 XX 04-OCT-2001; 2001WO-US031292.  
 XX  
 XX 20-MAR-2001; 2001US-00813432.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;  
 PI Vernet CAM;  
 XX  
 XX WPI; 2004-123380/12.  
 DR P-PSDB; ADI79323.  
 XX  
 XX Claim 8; Page 41; 158pp; English.  
 PS  
 CC The present invention relates to novel NOVX proteins and their coding  
 CC sequences (ADI79304-ADI79327). The sequences are useful for the  
 CC manufacture of a medicament for treating a syndrome associated with a  
 CC human disease associated with the protein, or for diagnosing and treating  
 CC disorders associated with the NOVX protein, such as cancer, myelogenous  
 CC leukaemia, congenital neonatal autoimmune thrombocytopaenia,  
 CC immunological disorders, allergy and infection, asthma, lung diseases,  
 CC reproductive disorders, male and female reproductive diseases,  
 CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,  
 CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,  
 CC hypertension, stroke and heart failure. NOV10a represents a new member of  
 CC the chemokine receptor family and the gene is located on chromosome 1.  
 CC  
 XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;  
 SQ  
 Query Match 100.0%; Score 1119; DB 12; Length 1119;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-306;  
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGAGCACACGACGACCCACCTCGCAGCCACAGCTCGTGTCTTGGTGTCCCGCGC 60  
 Db 1 ATGGAGCACACGACGACCCACCTCGCAGCCACAGCTCGTGTCTTGGTGTCCCGCGC 60  
 QY 61 TCGGCTCGGGCTTGGGTTTCGTGCGCCGTGTACTACAGCTCTTGTGCTCGGT 120  
 Db 61 TCGGCTCGGGCTTGGGTTTCGTGCGCCGTGTACTACAGCTCTTGTGCTCGGT 120  
 QY 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
 Db 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
 QY 181 TCCTCCTACAATCTCTTGGCACTCGTGTGCGGACATCTTGTCTCTTTTCATA 240  
 Db 181 TCCTCCTACAATCTCTTGGCACTCGTGTGCGGACATCTTGTCTCTTTTCATA 240  
 QY 241 GTGTTGTGGACTTCTCTGTGGAAGATTTTCATCTTGAACATGCAGATGCCTCAGTCCCC 300  
 Db 241 GTGTTGTGGACTTCTCTGTGGAAGATTTTCATCTTGAACATGCAGATGCCTCAGTCCCC 300

QY 301 GACAGATCATAGAGTGTCTGGATTCTCATCCATCCACACCTCCATATGGATTACTGTA 360  
 Db 301 GACAGATCATAGAGTGTCTGGATTCTCATCCATCCACACCTCCATATGGATTACTGTA 360  
 QY 361 CGGTTAAACCATTTGACAGGTATATCGTCTCTGCGACCCCGCTCAAGTACACACGGTCTCA 420  
 Db 361 CGGTTAAACCATTTGACAGGTATATCGTCTCTGCGACCCCGCTCAAGTACACACGGTCTCA 420  
 QY 421 TACCCAGCCCGACCCCGGAAAGTCAATTGTAAGTGTTCATCATCCTGCTCTCTGACGAGC 480  
 Db 421 TACCCAGCCCGACCCCGGAAAGTCAATTGTAAGTGTTCATCATCCTGCTCTCTGACGAGC 480  
 QY 481 ATCCCTATTACTGTGTGGCCCAACATCTGGAGTGNAGACTACATCAGCACCTCTGTGAT 540  
 Db 481 ATCCCTATTACTGTGTGGCCCAACATCTGGAGTGNAGACTACATCAGCACCTCTGTGAT 540  
 QY 541 CAGTCCTCATCTGGATCCACTGTCTTACCGCTTACCTGTGGTGGCCCTGCTCCATCTTCTTC 600  
 Db 541 CAGTCCTCATCTGGATCCACTGTCTTACCGCTTACCTGTGGTGGCCCTGCTCCATCTTCTTC 600  
 QY 601 ATCTTGAACATCAATCTGTGTACAGCTCAGGAGGAGCAATTTTCGTCTCCGTGGC 660  
 Db 601 ATCTTGAACATCAATCTGTGTGTACAGCTCAGGAGGAGCAATTTTCGTCTCCGTGGC 660  
 QY 661 TACTCCAGGGGAGAGACCCGCCATCTTGTTCACCATTACCTCCATCTTGGCCACACTT 720  
 Db 661 TACTCCAGGGGAGAGACCCGCCATCTTGTTCACCATTACCTCCATCTTGGCCACACTT 720  
 QY 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGGCCCATCCAGAACCGC 780  
 Db 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGGCCCATCCAGAACCGC 780  
 QY 781 TGGTGTGTACATCATGTCCGACATTCGCAATTCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
 Db 781 TGGTGTGTACATCATGTCCGACATTCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
 QY 841 AACTTCTTCTTACTGTCTTATCAGCAGCGGTTCGACACATGGGAGCGCCGACGCTC 900  
 Db 841 AACTTCTTCTTACTGTCTTATCAGCAGCGGTTCGACACATGGGAGCGCCGACGCTC 900  
 QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTCTCAGTTTCTACCAATCATATACTTTTCC 960  
 Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTCTCAGTTTCTACCAATCATATACTTTTCC 960  
 QY 961 ATAAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACATCGCATCAAGATCTGTGTGAC 1020  
 Db 961 ATAAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACATCGCATCAAGATCTGTGTGAC 1020  
 QY 1021 CAGTATGCAAAAATGGAAAACCTATATAAAGTGTATATGACAGAAAAGCTCTTACCAG 1080  
 Db 1021 CAGTATGCAAAAATGGAAAACCTATATAAAGTGTATATGACAGAAAAGCTCTTACCAG 1080  
 QY 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCATCTGTGA 1119  
 Db 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCATCTGTGA 1119  
 RESULT 4  
 ID ADO56001 standard; cDNA; 1119 BP.  
 XX  
 AC ADO56001;  
 XX  
 DT 15-JUL-2004 (first entry)  
 XX  
 XX DNA encoding human NOV10.  
 XX human; gene; ss; cancer; obesity; diabetic nephropathy;  
 KW acute pancreatitis; stroke; multiple sclerosis.  
 XX  
 OS Homo sapiens.  
 XX

PN US2004058862-A1.  
XX 25-MAR-2004.  
XX 18-SEP-2002; 2002US-00246583.  
XX 18-SEP-2002; 2002US-00246583.  
XX (MAJU//) MAJUMDER K.  
XX Majumder K;  
XX WPI; 2004-268835/25.  
XX P-PSDB; ADO56002.  
XX Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic  
XX nephropathy, acute pancreatitis, strokes and multiple sclerosis.  
XX Claim 9; Page 26-27; 87pp; English.  
XX The invention relates to novel isolated NOVX nucleic acids and encoded  
XX polypeptides. The nucleic acids, polypeptides and antibodies raised  
XX against the polypeptides are useful for preventing or treating diseases  
XX associated with aberrant NOVX expression or activity e.g., cancer,  
XX obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple  
XX sclerosis. The present sequence represents a NOVX nucleic acid of the  
XX invention.  
XX SQ Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;  
Query Match 100.0%; Score 1119; DB 12; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 1.3e-306;  
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACAGCAGCGCCACCTCGCAGCCAGCAAGCTCGCTGCTTGGTGGTCCCGCG 60  
DB 1 ATGGAGCACAGCAGCGCCACCTCGCAGCCAGCAAGCTCGCTGCTTGGTGGTCCCGCG 60  
QY 61 TCGGCTCGCGCTTGGGTTTGGTGGCGGCTGCTACTACAGCCTTGTGTGCTCGGT 120  
DB 61 TCGGCTCGCGCTTGGGTTTGGTGGCGGCTGCTACTACAGCCTTGTGTGCTCGGT 120  
QY 121 TTACAGCAAAATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
DB 121 TTACAGCAAAATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
QY 181 TCCTCTTACAACTATCTCTTGGCACTGCTGCTGCGGCACTTGTGCTCTTTTCATA 240  
DB 181 TCCTCTTACAACTATCTCTTGGCACTGCTGCTGCGGCACTTGTGCTCTTTTCATA 240  
QY 241 GTGTTTGTGACTTCTCTTGGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCG 300  
DB 241 GTGTTTGTGACTTCTCTTGGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCG 300  
QY 301 GACAAATCATAGAGTGTGGAATTTCTATCATCCACACCTCCATATGATGATCTGTA 360  
DB 301 GACAAATCATAGAGTGTGGAATTTCTATCATCCACACCTCCATATGATGATCTGTA 360  
QY 361 CCGTTAACCAATTCACAGGTATATCGCTGCTGCTGCGGCACTTGTGCTCTTTCATA 420  
DB 361 CCGTTAACCAATTCACAGGTATATCGCTGCTGCTGCGGCACTTGTGCTCTTTCATA 420  
QY 421 TACCCAGCCGCGACCCGGAAGTCAATTTGAAGTGTGTTTACATCACCTGCTTCTTGACGAG 480  
DB 421 TACCCAGCCGCGACCCGGAAGTCAATTTGAAGTGTGTTTACATCACCTGCTTCTTGACGAG 480  
QY 481 ATCCCTTATTTACCTGGTGGCCCAACATCTGGAATGGAAGATACATCAGCACTTGTGCAT 540  
DB 481 ATCCCTTATTTACCTGGTGGCCCAACATCTGGAATGGAAGATACATCAGCACTTGTGCAT 540  
QY 541 CAGGTCTCATCTGGATCCACTGCTTTCACGGTCTACCTGGTGGCTGCTCCATCTTCTTC 600  
DB 541 CAGGTCTCATCTGGATCCACTGCTTTCACGGTCTACCTGGTGGCTGCTCCATCTTCTTC 600

QY 601 ATCTTGAACCTCAATCAATTTGTGTACAGCTCAGGAGGAGAGCAATTTTCTCTCCGTGGC 660  
DB 601 ATCTTGAACCTCAATCAATTTGTGTACAGCTCAGGAGGAGAGCAATTTTCTCTCCGTGGC 660  
QY 661 TACTCCACGGGAGAGACACCGCCATCTTGTTCACCACTTACCTCCATCTTTGCCACATT 720  
DB 661 TACTCCACGGGAGAGACACCGCCATCTTGTTCACCACTTACCTCCATCTTTGCCACATT 720  
QY 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
DB 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
QY 781 TGGCTGGTACACATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 840  
DB 781 TGGCTGGTACACATCATGATTCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 840  
QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACATGCGCCACGCTC 900  
DB 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACATGCGCCACGCTC 900  
QY 901 AAGGCTTCTTCAAGTGCAGAGCAACTGTACAGTTCTTACACCAATCATTAATTTTCC 960  
DB 901 AAGGCTTCTTCAAGTGCAGAGCAACTGTACAGTTCTTACACCAATCATTAATTTTCC 960  
QY 961 ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACTGTCATCAAGATGCTGTGTATC 1020  
DB 961 ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACTGTCATCAAGATGCTGTGTATC 1020  
QY 1021 CAGTATGACAAAATGGAACCTATATAAAGTGTATATGACAGCAAAAGCTCTTACCAG 1080  
DB 1021 CAGTATGACAAAATGGAACCTATATAAAGTGTATATGACAGCAAAAGCTCTTACCAG 1080  
QY 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTCTGTGA 1119  
DB 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTCTGTGA 1119  
RESULT 5  
AAD29672  
ID AAD29672 standard; cDNA; 1130 BP.  
XX  
AC AAD29672;  
XX  
DT 17-MAY-2002 (first entry)  
XX  
DE Human G-protein coupled receptor (GCRC-6) cDNA.  
XX  
KW Human; G-protein coupled receptor; GCRC-6; cell proliferative disorder;  
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW anemias; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
KW neurotropic; neuroprotective; cardiac; immunosuppressive; anorectic;  
KW virucide; gene; SS.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 12..1130  
FT /tag= a  
FT /product= "Human GCRC-6"  
FT sig\_peptide 12..152  
FT /tag= b  
FT mat\_peptide 153..1127  
FT /tag= c  
FT /product= "Human mature GCRC-6"  
XX  
PN WO200210387-A2.  
XX

PD 07-FEB-2002.  
 XX 25-JUL-2001; 2001WO-US023433.  
 XX 27-JUL-2000; 2000US-0221478P.  
 PR 03-AUG-2000; 2000US-0233268P.  
 PR 21-AUG-2000; 2000US-0227054P.  
 PR 08-SEP-2000; 2000US-0231121P.  
 PR 13-SEP-2000; 2000US-0232243P.  
 PR 15-SEP-2000; 2000US-0232691P.  
 PR 22-SEP-2000; 2000US-0235146P.  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
 PI Elliot VS, Ramanujam J, Baughn MR, Kallick DA, Wallia NK, Hafalia AJA;  
 PI Yao MG, Lu Y, Tribouley CN, Policky JL, Kearney L, Gaul RC;  
 PI Warren BA, Lee EA, Ding L;  
 XX WPI: 2002-188744/24.  
 DR P-PSDB; AAE18645.  
 XX  
 PT New human G-protein coupled receptor polypeptide for diagnosis,  
 PT prevention and treatment of cell proliferative, neurological,  
 PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic  
 PT disorders.  
 XX  
 PS Claim 5; Page 141; 150pp; English.  
 XX  
 CC The invention relates to novel human G-protein coupled receptors (GPCR)  
 CC and their encoding polynucleotides. GPCR is useful as an immunogen for  
 CC preparing monoclonal and polyclonal antibodies. GPCR is useful for  
 CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,  
 CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,  
 CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's  
 CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,  
 CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,  
 CC cirrhosis, Crohn's disease), an autoimmune/inflammatory disorder (e.g.,  
 CC acquired immunodeficiency syndrome (AIDS), allergy, anemia, asthma,  
 CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,  
 CC osteoporosis), and viral infections. GPCR is useful in a number of drug  
 CC screening techniques, and to analyse the proteome of a tissue or cell  
 CC type. GPCR is useful for creating knockin humanised animals or  
 CC transgenic animals to model human diseases, in somatic or germline gene  
 CC therapy, to generate a transcript image of a tissue or cell type, for  
 CC detecting differences in the chromosomal location due to translocation,  
 CC inversion, etc., among normal, carrier or affected individuals, and as  
 CC hybridization probes for mapping naturally occurring genomic sequences.  
 CC GPCR is useful in Southern or northern analysis, dot blot or other  
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,  
 CC multifunctional enzyme linked immunosorbent (ELISA)-like assays, and in  
 CC microarrays utilising fluids or tissues from patients to detect altered  
 CC GPCR expression. The present sequence is human GPCR-6 cDNA  
 XX  
 SQ Sequence 1130 BP; 258 A; 357 C; 226 G; 289 T; 0 U; 0 Other;  
 Query Match 100.0%; Score 1119; DB 6; Length 1130;  
 Best Local Similarity 100.0%; Pred. NO. 1.3e-306;  
 Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGAGCACAGCAGCCCACTCGCAGCAACAGCTCGCTGTTGGTGGTCCCGCGC 60  
 DB 1119  
 QY 61 TCGGCTCGCGCTTGGGTTTCGTGGCCGTTGCTACTACAGCCTCTTGTGTCCTCGGT 120  
 DB 72 TCGGCTCGCGCTTGGGTTTCGTGGCCGTTGCTACTACAGCCTCTTGTGTCCTCGGT 131  
 QY 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180  
 DB 132 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 191  
 QY 181 TCCTCTACAACTATCTCTTGGCACTCGCTGTCGCGACATCTTGTCTCTTTTCATA 240

DB 192 TCCTCTACAACTATCTCTTGGCACTCGCTGTCGCGACATCTTGTCTCTTTTCATA 251  
 QY 241 GTGTTTGGGACTTCTCTGTTGGAAGATTTCATCTTTGAAATGCGAGATGCTCAGGTCCCC 300  
 DB 252 GTGTTTGGGACTTCTCTGTTGGAAGATTTCATCTTTGAAATGCGAGATGCTCAGGTCCCC 311  
 QY 301 GACAGATCATAGAGTCTGGAAATTTCTATCCATCCACACCTCCATATGGATTACTGTA 360  
 DB 312 GACAAAGATCATAGAGTCTGGAAATTTCTATCCATCCACACCTCCATATGGATTACTGTA 371  
 QY 361 CGTTTAAACATTGACAGGTATATCGCTGCTGCGCACCCGCTCAAGTACACACGCTCTCA 420  
 DB 372 CGTTTAAACATTGACAGGTATATCGCTGCTGCGCACCCGCTCAAGTACACACGCTCTCA 431  
 QY 421 TACCAGCCCCGACCCCGAAAGTCAATTAAGTGTTTACATCACTGCTCTCTGACACGAGC 480  
 DB 432 TACCAGCCCCGACCCCGAAAGTCAATTAAGTGTTTACATCACTGCTCTCTGACACGAGC 491  
 QY 481 ATCCCTTATTACTGTTGGGCCCCAATCTGGGACTGGAAGTACATCAGCACCTCTGTGCAT 540  
 DB 492 ATCCCTTATTACTGTTGGGCCCCAATCTGGGACTGGAAGTACATCAGCACCTCTGTGCAT 551  
 QY 541 CAGGTCTCTCATCTGGATCCACTGCTTACCGCTTACCTGGTGGCCCTGCTCCATCTTCTTC 600  
 DB 552 CAGGTCTCTCATCTGGATCCACTGCTTACCGCTTACCTGGTGGCCCTGCTCCATCTTCTTC 611  
 QY 601 ATCTTGAATCAATCATCTGTTGTAAGCTCAGGAGGAAGAGCAATTTTGGTCTCGTGGC 660  
 DB 612 ATCTTGAATCAATCATCTGTTGTAAGCTCAGGAGGAAGAGCAATTTTGGTCTCGTGGC 671  
 QY 661 TACTCCAGGGGAGACACCGCCATCTGTTTACCATTTACCATTTCTTGGCCACACTT 720  
 DB 672 TACTCCAGGGGAGACACCGCCATCTGTTTACCATTTACCATTTCTTGGCCACACTT 731  
 QY 721 TGGGCCCCCGCATCATCATGATCTTTTACCACTCTTATGGGGGCCCATCCAGAACCGC 780  
 DB 732 TGGGCCCCCGCATCATCATGATCTTTTACCACTCTTATGGGGGCCCATCCAGAACCGC 791  
 QY 781 TGGTGGTACATCATCATGTCGACATTCGCCAATTCGCAATGTAGCCCTTTCTGAACACAGCCATC 840  
 DB 792 TGGTGGTACATCATCATGTCGACATTCGCCAATTCGCAATGTAGCCCTTTCTGAACACAGCCATC 851  
 QY 841 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTTCCGACCATGGCAGCGCCACGCTC 900  
 DB 852 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTTCCGACCATGGCAGCGCCACGCTC 911  
 QY 901 AAGGCTTTCTTCAAGTGCAGAACCTGTACAGTTTCTACCAATTCATAAATCTTTTCC 960  
 DB 912 AAGGCTTTCTTCAAGTGCAGAACCTGTACAGTTTCTACCAATTCATAAATCTTTTCC 971  
 QY 961 ATAAACAGTACCCCTGGATCTCGCGGCAAACTACATGCTGATCAAGATGCTGGGTGAC 1020  
 DB 972 ATAAACAGTACCCCTGGATCTCGCGGCAAACTACATGCTGATCAAGATGCTGGGTGAC 1031  
 QY 1021 CAGTATGCAAAATGGAACCTTATAAAAGTGTATGACAGCAAAAGCTCTCTACAG 1080  
 DB 1032 CAGTATGCAAAATGGAACCTTATAAAAGTGTATGACAGCAAAAGCTCTCTACAG 1091  
 QY 1081 TTTGAAGATGCCATTGGAGCTTGTGTCTATCATCTCTGTGA 1119  
 DB 1092 TTTGAAGATGCCATTGGAGCTTGTGTCTATCATCTCTGTGA 1130  
 RESULT 6  
 ID ACF05275 standard; cDNA; 2198 BP.  
 XX ACF05275;  
 AC ACF05275;  
 XX 06-NOV-2003 (first entry)  
 XX Human G-protein coupled receptor HGPBMY34 cDNA.





RESULT 7  
ACF05276  
ID ACF05276 standard; cDNA; 1110 BP.  
XX AC ACF05276;  
XX 06-NOV-2003 (first entry)  
XX  
DE Human G-protein coupled receptor HGPBMY34 variant cDNA.  
XX  
XX HGPBMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;  
KW human; neuroprotective; nootropic; tranquilizer; antimigraine;  
KW neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;  
KW cytostatic; cardiant; hypotensive; antianginal; analgesic; anorectic;  
KW anti-HIV; antiasthmatic; osteopathic; uropathic; antiulcer; antiallergic;  
KW gene therapy; gene; ss.  
XX  
XX Homo sapiens.  
XX  
XX  
XX Key Location/Qualifiers  
FT CDS 1..1110  
FT /\*tag= a  
FT /product= "HGPBMY34 variant"  
XX  
XX WO2003050256-A2.  
XX  
XX 19-JUN-2003.  
XX  
XX 06-DEC-2002; 2002WO-US039290.  
XX  
XX 06-DEC-2001; 2001US-0338371P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
XX Feder JN, Gopal S, Mintier GA, Ramanathan CS;  
XX  
XX WPI: 2003-577295/54.  
XX P-PSDB; ABR62522.  
XX  
XX New nucleic acid molecule encoding a human G-protein coupled receptor,  
PT HGPBMY34, useful for diagnosing, preventing or treating diseases  
PT involving the receptor, for example Parkinson's disease, dementia,  
PT asthma, hypertension or cancer.  
XX  
XX Claim 1; Fig 2A-B; 112pp; English.  
XX  
XX The present sequence is that of cDNA encoding a variant of human  
CC HGPBMY34 containing a deletion of 3 amino acids of the protein sequence  
CC but expected to share at least some of the expression patterns and  
CC function of HGPBMY34. This is a newly identified G-protein coupled  
CC receptor (GPCR) belonging to the group of 'Class A' GPCRs and showing  
CC homology to the Pfam model 7 transmembrane receptor, rhodopsin family.  
CC HGPBMY34 (or GPCR-P14, GPCR-145) is highly expressed in brain (amygdala,  
CC caudate nucleus, corpus callosum, hippocampus, thalamus, substantia  
CC nigra), spinal cord and pituitary, indicating an association in  
CC neurological systems and conditions. It is also expressed in the bone  
CC marrow and testis. The invention provides HGPBMY34 polynucleotides,  
CC polypeptides and antibodies, expression vectors, host cells and antisense  
CC molecules, methods for screening for modulators of HGPBMY34 activity  
CC and/or function, and methods for diagnosing, treating, preventing and  
CC screening for disorders and diseases associated with abnormal HGPBMY34  
CC activity, including: a disorder related to aberrant G-protein coupled  
CC signaling; a disorder related to aberrant cell cycle regulation;  
CC neurological disorders; anxiety; headache; migraine; schizophrenia; manic  
CC depression; delirium; dementia; severe mental retardation and  
CC dyskinesias, such as Huntington's disease or Gilles de la Tourette's  
CC syndrome; Parkinson's disease; brain disorders; spinal cord disorders;  
CC affective disorders; neoplastic disorders; cardiovascular disorders;  
CC acute heart failure; hypotension; hypertension; angina pectoris;  
CC myocardial infarction; an immunological disorder; immune-related  
CC disorders; endocrinal diseases; growth disorders; neuropathic pain;

CC obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis  
CC ; psychosis; metabolic disorders; pituitary disorders; urinary retention;  
CC ulcers; allergies; or benign prostatic hypertrophy (all claimed)

XX SQ Sequence 1110 BP; 253 A; 351 C; 220 G; 286 T; 0 U; 0 Other;  
Query Match 97.5%; Score 1091; DB 9; Length 1110;  
Beat Local Similarity 99.2%; Pred. No. 1.2e-298;  
Matches 1110; Conservative 0; Mismatches 0; Indels 9; Gaps 1;  
QY 1 ATGAGACACACGACGCCCCACCTCGACGCCAACAGCTCGTGTCTGTGTGTCGCCCGCC 60  
DB 1 ATGAGACACACGACGCCCCACCTCGACGCCAACAGCTCGTGTCTGTGTGTCGCCCGCC 60  
QY 61 TCGGCTTGGGCTTGGGTTTGGTCCCGTGTCTTACTACAGCCTTGTCTGTCTCGCT 120  
DB 61 TCGGCTTGGGCTTGGGTTTGGTCCCGTGTCTTACTACAGCCTTGTCTGTCTCGCT 120  
QY 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCACAGCTGGTGGCAAGACAGAAAG 180  
DB 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCACAGCTGGTGGCAAGACAGAAAG 180  
QY 181 TCCTCTACAACTATCTCTTGGCACTCGCTGTGCGACATCTTGGTCTCTTTTTCATA 240  
DB 181 TCCTCTACAACTATCTCTTGGCACTCGCTGTGCGACATCTTGGTCTCTTTTTCATA 240  
QY 241 GTGTTTGGGACTTCTCTTGGAGATTTCATCTTGAACATGCAGATGCCTCAGTCTCCC 300  
DB 241 GTGTTTGGGACTTCTCTTGGAGATTTCATCTTGAACATGCAGATGCCTCAGTCTCCC 300  
QY 301 GACAAAGATCATAGAAAGTCTGGAATTCATCTCCATCCACACCTCCATATGATTACTGTA 360  
DB 301 GACAAAGATCATAGAAAGTCTGGAATTCATCTCCATCCACACCTCCATATGATTACTGTA 360  
QY 361 CCGTTAACCAATTGACAGGTATATCGTGTCTGCCACCCGCTCAAGTACCAACACGCTCTCA 420  
DB 361 CCGTTAACCAATTGACAGGTATATCGTGTCTGCCACCCGCTCAAGTACCAACACGCTCTCA 420  
QY 421 TACCAGCCGCGACCCGGAAGTCAATTGTAAGTGTTTACATCACCTGCTTCTGACGACG 480  
DB 421 TACCAGCCGCGACCCGGAAGTCAATTGTAAGTGTTTACATCACCTGCTTCTGACGACG 480  
QY 481 ATCCCTATTACTGTGTGCCCCAAATCTGGAAGTGAAGTACATCAGCACCTCTGTGCA 540  
DB 481 ATCCCTATTACTGTGTGCCCCAAATCTGGAAGTGAAGTACATCAGCACCTCTGTGCA 540  
QY 541 CAGTCTCATCTGGATCCACTGTCTTACCGCTTACCTGTGCGCTGTCTCATCTTCTTC 600  
DB 541 CAGTCTCATCTGGATCCACTGTCTTACCGCTTACCTGTGCGCTGTCTCATCTTCTTC 600  
QY 601 ATCTTGAATCAATCTGTGTACAGCTCAGGAGGAGAGCAATTTTCGTCTCGTGGC 660  
DB 601 ATCTTGAATCAATCTGTGTACAGCTCAGGAGGAGAGCAATTTTCGTCTCGTGGC 660  
QY 661 TACTCCAGGGAAGACCCCGCATCTTGTTCACCATTTACTCTCATCTTTGCCACACTT 720  
DB 661 TACTCCAGGGAAGACCCCGCATCTTGTTCACCATTTACTCTCATCTTTGCCACACTT 720  
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTTACACCTCTTATGGGGCGCCCATCCAGAACCG 780  
DB 721 TGGGCCCCCGCATCATCATGATTTCTTTTACACCTCTTATGGGGCGCCCATCCAGAACCG 780  
QY 781 TGGTGTGTACACATCATGTCTCGACATTTGCCAACATGTAGCCCTTCTGAAACAGCCATC 840  
DB 781 TGGTGTGTACACATCATGTCTCGACATTTGCCAACATGTAGCCCTTCTGAAACAGCCATC 840  
QY 841 AACTTCTTCTTCTACTGTCTTCATCAGCAAGCGGTTCCGACATGCGGAGCGCCACGCTC 900  
DB 841 AACTTCTTCTTCTACTGTCTTCATCAGCAAGCGGTTCCGACATGCGGAGCGCCACGCTC 900  
QY 901 AAGGCTTTCTTCAAGTGTCCAGAACCACTGTACAGTTTCTACACCAATCAATCTTTTCC 960  
DB 901 AAGGCTTTCTTCAAGTGTCCAGAACCACTGTACAGTTTCTACACCAATCAATCTTTTCC 960

QY 961 ATAAAGTAGCCCTGGATCTCGCGGCAAACTCACTGCATCAAGATGCTGGGTAC 1020  
| | | | |  
Db 961 ATAAAGTAGCCCTGGATCTCGCGGCAAACTCACTGCATCAAGATGCTGGGTAC 1020  
| | | | |  
QY 1021 CAGTATGACAAAATGGAACCTATATAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 1080  
| | | | |  
Db 1021 CAGTATGACAAAATG-----AAAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 1071  
| | | | |

QY 1081 TTTGAAGATGCCATTGGAGCTTGTCATCATCTCTGTA 1119  
| | | | |  
Db 1072 TTTGAAGATGCCATTGGAGCTTGTCATCATCTCTGTA 1110  
| | | | |

RESULT 8  
ADD18120  
ID ADD18120 standard; DNA; 2189 BP.

AC ADD18120;

XX 15-JAN-2004 (first entry)

DT Human G-protein coupled receptor (GPCR) DNA sequence Seq ID6.

XX G protein coupled receptor; GPCR; signal transduction pathway; G protein;  
KW Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;  
KW colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;  
KW sleeplessness; hypertension; anxiety; stress; renal failure;  
KW cardiovascular disorder; neural disorder; oncology disorder;  
KW immune disorder; neuroprotective; gene therapy; gene; ds.

XX Homo sapiens.

XX Key Location/Qualifiers  
FH 1047..-2157  
FT CDS

FT /tag= a  
FT /product= "Human G-protein coupled receptor"  
FT /transl\_except= (2085..2091, aa:KPIKS)

XX WO2003016478-A2.

XX 27-FEB-2003.

XX 15-AUG-2002; 2002WO-US026017.

XX 20-AUG-2001; 2001US-0313658P.

XX 12-SEP-2001; 2001US-0318675P.

XX 30-OCT-2001; 2001US-0340703P.

XX 26-NOV-2001; 2001US-0333417P.

XX 06-DEC-2001; 2001US-0338367P.

XX 06-FEB-2002; 2002US-0355596P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Feder JN, Ramanathan CS, Gopal S, Mintier GA;

XX WPI; 2003-278558/27.

XX P-PSDB; ADD18023.

XX New nucleic acid, useful for manufacturing a medicament for preventing,  
XX treating or ameliorating a medical condition e.g., neural disorder.

XX Example 2; SEQ ID NO 6; 251pp; English.

XX This invention relates to novel G protein coupled receptors (GPCRs) and  
XX their encoding nucleotide sequences. Many medically significant  
XX biological processes are mediated by proteins participating in signal  
XX transduction pathways involving G proteins. GPCRs are one of the largest  
XX receptor superfamilies known. These receptors are biologically important  
XX and malfunction of these receptors results in diseases such as  
XX Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal  
XX pigmentosa and asthma. They are also involved in depression,  
XX schizophrenia, sleeplessness, hypertension, anxiety, stress, renal

CC failure and other cardiovascular, neural, oncology and immune disorders.  
CC A modulator of the GPCRs of the invention may have neuroprotective  
CC activity whilst the sequences of the invention may be useful for gene  
CC therapy. The invention may also be useful for manufacturing a medicament  
CC for preventing, treating or ameliorating a medical condition. The present  
CC sequence is the DNA sequence encoding a human GPCR of the invention.

XX SQ Sequence 2189 BP; 451 A; 691 C; 581 G; 466 T; 0 U; 0 Other;

Query Match 97.5%; Score 1091; DB 10; Length 2189;  
Best Local Similarity 99.2%; Pred. No. 1.6e-298;  
Matches 1110; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY 1 ATGGAGCACACGACGCCACCTCGCAGCCAAAGCTGCTGTCTTGGTGGTCCCGCGC 60  
| | | | |  
Db 1047 ATGGAGCACACGACGCCACCTCGCAGCCAAAGCTGCTGTCTTGGTGGTCCCGCGC 1106  
| | | | |

QY 61 TCGGGCTGGGCTTGGTTCGTCGCCGTGCTACTACAGCCTCTTGTGTGCTCGCT 120  
| | | | |  
Db 1107 TCGGGCTGGGCTTGGTTCGTCGCCGTGCTACTACAGCCTCTTGTGTGCTCGCT 1166  
| | | | |

QY 121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 180  
| | | | |  
Db 1167 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 1226  
| | | | |

QY 181 TCCTCTCAAACTATCTCTTGGCACTCGTCTGCGGACATCTTGGTCTCTTTTTCATA 240  
| | | | |  
Db 1227 TCCTCTCAAACTATCTCTTGGCACTCGTCTGCGGACATCTTGGTCTCTTTTTCATA 1286  
| | | | |

QY 241 GTGTTTGGACTTCTGTGTGGAAGATTTCATCTTGAACATGACAGTGGCTCAGGTCCCC 300  
| | | | |  
Db 1287 GTGTTTGGACTTCTGTGTGGAAGATTTCATCTTGAACATGACAGTGGCTCAGGTCCCC 1346  
| | | | |

QY 301 GACAAAGATCATAGAAGTGTCTGGAATTCATCATCCACACCTCCATATGGAATTACTGTA 360  
| | | | |  
Db 1347 GACAAAGATCATAGAAGTGTCTGGAATTCATCATCCACACCTCCATATGGAATTACTGTA 1406  
| | | | |

QY 361 CGGTTAACCATTTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACCACAGGCTCTCA 420  
| | | | |  
Db 1407 CGGTTAACCATTTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACCACAGGCTCTCA 1466  
| | | | |

QY 421 TACCAGCCCGCACCCGGAAAGTCATTGTAAAGTGTATACATCACCTGCTTCTTGACCAGC 480  
| | | | |  
Db 1467 TACCAGCCCGCACCCGGAAAGTCATTGTAAAGTGTATACATCACCTGCTTCTTGACCAGC 1526  
| | | | |

QY 481 ATCCCTCTATTAGTGTGGCCCAACATCTGAGCTGAAGACTACATCAGACACCTCTGTGCAAT 540  
| | | | |  
Db 1527 ATCCCTCTATTAGTGTGGCCCAACATCTGAGCTGAAGACTACATCAGACACCTCTGTGCAAT 1586  
| | | | |

QY 541 CAGGTCTCTCATCTGGATCCACTGTCTTCAACCGTCTACCTGGTGCCTGCTCCATCTTCTTC 600  
| | | | |  
Db 1587 CAGGTCTCTCATCTGGATCCACTGTCTTCAACCGTCTACCTGGTGCCTGCTCCATCTTCTTC 1646  
| | | | |

QY 601 ATCTTGAATCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTGTCTCCGTGGC 660  
| | | | |  
Db 1647 ATCTTGAATCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTGTCTCCGTGGC 1706  
| | | | |

QY 661 TACTCCAGGGGAAGACACCGCCCATCTTGTTCACCATTTACCTTCCATCTTTGCGACACTT 720  
| | | | |  
Db 1707 TACTCCAGGGGAAGACACCGCCCATCTTGTTCACCATTTACCTTCCATCTTTGCGACACTT 1766  
| | | | |

QY 721 TGGGGCCCCCGGCATCATGATTTTACACACCTCTTATGGGGCGGCCCATCCAGAAACCGC 780  
| | | | |  
Db 1767 TGGGGCCCCCGGCATCATGATTTTACACACCTCTTATGGGGCGGCCCATCCAGAAACCGC 1826  
| | | | |

QY 781 TGGCTGGTACACATCATGTCCGACATTTGCAACATGCTAGCCCTTCTGAAACAGAGCATC 840  
| | | | |  
Db 1827 TGGCTGGTACACATCATGTCCGACATTTGCAACATGCTAGCCCTTCTGAAACAGAGCATC 1886  
| | | | |

QY 841 AACTCTTCTCTACTGCTTTCATCAGCAGCGGTTCCGACCATGGCAGCCGCGCCAGCTC 900  
| | | | |  
Db 1887 AACTCTTCTCTACTGCTTTCATCAGCAGCGGTTCCGACCATGGCAGCCGCGCCAGCTC 1946  
| | | | |

QY 901 AAGGCTTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTTACACCAATCATAACTTTCC 960  
Db |||||||  
QY 1947 AAGGCTTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTTACACCAATCATAACTTTCC 2006  
Db |||||||  
QY 961 ATAACAGTAGCCCTTGGATCTCGCGGCAAACTCACATGATCAAGATGCTGTGTAC 1020  
Db |||||||  
QY 2007 ATAACAGTAGCCCTTGGATCTCGCGGCAAACTCACATGATCAAGATGCTGTGTAC 2066  
QY 1021 CAGTATGACAAAATGGAACCTATATAAAGTCTGTATGACAGCAAAAGCTCTACACG 1080  
Db |||||||  
QY 2067 CAGTATGACAAAATGGAACCTATATAAAGTCTGTATGACAGCAAAAGCTCTACACG 2117  
QY 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGA 1119  
Db |||||||  
QY 2118 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGA 2156

RESULT 9  
ABN84269  
ID ABN84269 standard; cDNA; 1059 BP.  
XX  
AC ABN84269;  
XX  
DT 23-SEP-2002 (first entry)  
XX  
DE Human chemokine-like receptor cDNA.  
XX  
KW Chemokine-like receptor; G-protein coupled receptor; receptor; human;  
KW HIV infection; cardiovascular disease; asthma;  
KW chronic obstructive pulmonary disease; cardiant; antiasthmatic;  
KW vasotropic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;  
KW antiinflammatory; antiallergic; immunomodulator; gene therapy; gene; ss.  
XX  
OS Homo sapiens.

Key Location/Qualifiers  
FH 1. .1059  
FT CDS /tag= a  
FT /partial  
FT /product= "Chemokine-like receptor"  
FT /note= "the CDS does not include a stop codon"

XX  
PN WO200248358-A2.  
XX  
PD 20-JUN-2002.  
XX  
PF 12-DEC-2001; 2001WO-BP014571.  
XX  
PR 14-DEC-2000; 2000US-0255150P.  
PR 02-APR-2001; 2001US-0280110P.  
PR 21-JUN-2001; 2001US-0299474P.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;  
XX  
DR WPI; 2002-547858/58.  
DR P-PSDB; AB579517.  
XX  
PT New isolated polynucleotide encoding a chemokine-like receptor  
PT polypeptide for treating e.g. asthma, myocardial infarction, human  
PT immunodeficiency virus infection, or chronic obstructive pulmonary  
PT disease.  
XX  
PS Claim 1; Fig 1; 114pp; English.  
XX  
CC The present sequence is that of cDNA encoding a novel human chemokine-  
CC like receptor (see AN79517) of 353 amino acids and having 7 putative  
CC transmembrane domains, consistent with the structure of a G-protein  
CC coupled receptor. Its closest human homologue is C-C chemokine receptor  
CC 3. The novel receptor is expressed at low levels in most tissues. It is  
CC expressed at a high level in phytohaemagglutinin-stimulated CD8+ cells,  
CC but in none of the other immune cells tested. It may act as a receptor of

CC chemoattractant molecules on activated lymphocytes and be involved in  
CC cell trafficking and homing to sites of infection, inflammation or tissue  
CC injury. Regulation of activity of the novel receptor can therefore be  
CC used to treat cardiovascular, immunological and inflammatory diseases,  
CC including asthma and chronic obstructive pulmonary disease (COPD). The  
CC receptor may also be a target for viruses that reside in the nervous  
CC system. Regulating the binding of ligands, e.g. chemoattractant molecules  
CC or virus particles, to the receptor can therefore be used to modulate the  
CC immune response to inhibit viral infections, including HIV infection. A  
CC claimed method of reducing activity of the receptor involves contacting a  
CC cell with a reagent (preferably an antibody, antisense oligonucleotide or  
CC ribozyme) to a product (preferably RNA or a polypeptide) encoded by a  
CC polynucleotide encoding the human chemokine-like receptor in vivo or in  
CC vitro. A claimed method of treating a chemokine-like receptor dysfunction  
CC related disease selected from HIV infection, a cardiovascular disorder,  
CC asthma or COPD uses a reagent that modulates a function of the receptor  
XX  
SQ Sequence 1059 BP; 240 A; 342 C; 206 G; 271 T; 0 U; 0 Other;  
  
Query Match 93.8%; Score 1050; DB 6; Length 1059;  
Best Local Similarity 100.0%; Pred. No. 4.9e-287;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 ATGGAGCACACGACGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGTGCCCGCGC 60  
Db |||||||  
QY 61 TCGGCTCGCGCTTGGGTTTCGTCGCCGTGTCTACTACAGCCTTTCTGTGCTCGGT 120  
Db |||||||  
QY 61 TCGGCTCGCGCTTGGGTTTCGTCGCCGTGTCTACTACAGCCTTTCTGTGCTCGGT 120  
QY 121 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
Db |||||||  
QY 121 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
QY 181 TCCTCCTCAACTATCTCTTGGCACTGCTGCTGCGACATCTTGGTCTCTTTTTCATA 240  
Db |||||||  
QY 181 TCCTCCTCAACTATCTCTTGGCACTGCTGCTGCGACATCTTGGTCTCTTTTTCATA 240  
QY 241 GTGTTGTGGACTTCTGTTTGAAGATTTCACTTTGAACATGCGATGCGCTCAGGTCCCC 300  
Db |||||||  
QY 241 GTGTTGTGGACTTCTGTTTGAAGATTTCACTTTGAACATGCGATGCGCTCAGGTCCCC 300  
QY 301 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACACCTCCATATGATTAATGTA 360  
Db |||||||  
QY 301 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACACCTCCATATGATTAATGTA 360  
QY 361 CCGTTAACCATTTGACAGGTATATCGTGTCTGCGACATCTTGAACATGCGATGCGCTCAG 420  
Db |||||||  
QY 361 CCGTTAACCATTTGACAGGTATATCGTGTCTGCGACATCTTGAACATGCGATGCGCTCAG 420  
QY 421 TACCCAGCCCGCACCCCGGAAAGTCAATTGTAAGTGTGTTTACATCAGCTGCTTCTGAC 480  
Db |||||||  
QY 421 TACCCAGCCCGCACCCCGGAAAGTCAATTGTAAGTGTGTTTACATCAGCTGCTTCTGAC 480  
QY 481 ATCCCTTATTAATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540  
Db |||||||  
QY 481 ATCCCTTATTAATGCTGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTC 540  
QY 541 CAGTGTCTCATCTGATTCCTGCTTCACTGCTTACCTGTCGTCGTCGTCGTCGTCGTCGTC 600  
Db |||||||  
QY 541 CAGTGTCTCATCTGATTCCTGCTTCACTGCTTACCTGTCGTCGTCGTCGTCGTCGTCGTC 600  
QY 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAGAGCAATTTTCGTCCTCGTGGC 660  
Db |||||||  
QY 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAGAGCAATTTTCGTCCTCGTGGC 660  
QY 661 TACTTCCAGGGGAGAGACACCGCCATCTTGTTCACCATTAATCTTTCACACTTTTGGC 720  
Db |||||||  
QY 661 TACTTCCAGGGGAGAGACACCGCCATCTTGTTCACCATTAATCTTTCACACTTTTGGC 720  
QY 721 TGGGCCCCCCCCCATCATCATGATTTTACCATCTTATGGGGCGCCCATCCAGAACCGC 780  
Db |||||||

Db 721 TGGGCCCCCGCATCATCATGATCTTTTACCACCTCTATGGGGGCGCCCATCCAGAACCGC 780  
Qy 781 TGCGTGGTACATCATGTCGACATTCGCAATGCTAGCCCTTCTGAAACAGCCATC 840  
Db 781 TGCGTGGTACATCATGTCGACATTCGCAATGCTAGCCCTTCTGAAACAGCCATC 840  
Qy 841 AACTTCTTCTCTACTGCTTCTATCAGCAGCGGTTCCGACCATGGCAGCGCCAGCTC 900  
Db 841 AACTTCTTCTCTACTGCTTCTATCAGCAGCGGTTCCGACCATGGCAGCGCCAGCTC 900  
Qy 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC 960  
Qy 961 ATAACAAGTAGCCCTTGGATCTGCGCGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1020  
Db 961 ATAACAAGTAGCCCTTGGATCTGCGCGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1020  
Qy 1021 CAGTATGACAAAANTGGAAAACCTATAAA 1050  
Db 1021 CAGTATGACAAAANTGGAAAACCTATAAA 1050

## RESULT 10

ABK15562  
ID ABK15562 standard; cDNA; 1062 BP.  
XX  
AC ABK15562;  
XX  
DT 08-MAY-2002 (first entry)  
XX  
DE cDNA encoding novel G-protein coupled receptor TGR8, version #1.  
XX  
KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;  
KW cytosolic; antiinflammatory; antiulcer; fetal brain;  
KW central nervous system disease; circulatory organ disorder; cancer;  
KW metabolic disease; immunological disease; gastrointestinal disease;  
KW gene therapy; transgenic animal; human; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 1..1062  
FT CDS /\*tag= a  
FT /product= "TGR8"  
FT /note= "G-protein coupled receptor"  
XX  
PN WO200194582-A1.  
XX  
PD 13-DEC-2001.  
XX  
PF 01-JUN-2001; 2001WO-JP004643.  
XX  
PR 02-JUN-2000; 2000JP-00170446.  
XX 23-JUN-2000; 2000JP-00194926.  
XX (TAKA ) TAKEDA CHEM IND LTD.  
XX PA  
XX PI Terao Y, Matsui H, Shintani Y;  
XX WPI; 2002-164317/21.  
DR P-PSDB; AAU76416.  
XX  
PT Human fetal brain-originated G protein-coupled receptor protein TGR8 and  
PT encoding DNA, for developing drugs to treat e.g. diseases of the central  
PT nervous system or circulatory organs, cancer, and metabolic diseases.  
XX  
PS Claim 5; Page 94; 102pp; Japanese.  
XX  
CC The invention describes a human fetal brain-originated G protein-coupled  
CC receptor protein, or its salt. The protein and encoded DNA are useful for  
CC developing drugs to treat e.g. diseases of the central nervous system or  
CC circulatory organs, cancer, metabolic diseases, immunological diseases

CC and gastrointestinal diseases. The invention also describes creation of a  
CC probe for gene therapy and construction of a transgenic animal. This  
CC sequence encodes the novel G-protein coupled receptor TGR8, described in  
CC the method of the invention

SQ Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;

Query Match 93.8%; Score 1050; DB 6; Length 1062;  
Best Local Similarity 100.0%; Pred. No. 4.9e-287;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGAGCAGCAGCAGCCACCTGCGAGCCAGCAGCTGCTGTCTGTGGTGGTCCCGCGC 60  
Db 1 ATGGAGCAGCAGCAGCCACCTGCGAGCCAGCAGCTGCTGTCTGTGGTGGTCCCGCGC 60  
Qy 61 TCGGCTCGGGCTTGGGTTTGGTCCCGTGGTGTACTACAGCCTCTTGTGTGGTCTCGGT 120  
Db 61 TCGGCTCGGGCTTGGGTTTGGTCCCGTGGTGTACTACAGCCTCTTGTGTGGTCTCGGT 120  
Qy 121 TTACAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGAGAGAGAG 180  
Db 121 TTACAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGAGAGAGAG 180  
Qy 181 TCCTCTACAACTATCTTGGCACTGCTGCGGACATCTTGGTCTCTTTTTCATA 240  
Db 181 TCCTCTACAACTATCTTGGCACTGCTGCGGACATCTTGGTCTCTTTTTCATA 240  
Qy 241 GTGTTTGGGACTTCTCTGTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCG 300  
Db 241 GTGTTTGGGACTTCTCTGTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCG 300  
Qy 301 GACAAAGATCATAGAGTCTGGAATTTTCATCCATCCACACCTCCATATGGATTAATGTA 360  
Db 301 GACAAAGATCATAGAGTCTGGAATTTTCATCCATCCACACCTCCATATGGATTAATGTA 360  
Qy 361 CGGTTAAACATGACAGGTATATCGTCTGTGCGACCCGCTCAAGTACCAACAGGTCTCA 420  
Db 361 CGGTTAAACATGACAGGTATATCGTCTGTGCGACCCGCTCAAGTACCAACAGGTCTCA 420  
Qy 421 TACCAGCCCGCAGCCGGAAGTCAATGTAAGTGTATACATCACCTGCTCTCTGACCAGC 480  
Db 421 TACCAGCCCGCAGCCGGAAGTCAATGTAAGTGTATACATCACCTGCTCTCTGACCAGC 480  
Qy 481 ATCCCTTATTAATGTTGGTGGCCCAACATCTGGAATGAAGACTACATCAGACCTCTGTCAT 540  
Db 481 ATCCCTTATTAATGTTGGTGGCCCAACATCTGGAATGAAGACTACATCAGACCTCTGTCAT 540  
Qy 541 CAGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGCCCTGCTCCATCTCTTTC 600  
Db 541 CAGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGCCCTGCTCCATCTCTTTC 600  
Qy 601 ATCTTGAATCAATCATTTGTGTACAGCTCAGGAGAGAGCAATTTTCTGCTCCGTGGC 660  
Db 601 ATCTTGAATCAATCATTTGTGTACAGCTCAGGAGAGAGCAATTTTCTGCTCCGTGGC 660  
Qy 661 TACTCCAGGGGAGAGACCCAGCCATCTTGTTCACCAATTTACCTCCATCTTTGSCACATTT 720  
Db 661 TACTCCAGGGGAGAGACCCAGCCATCTTGTTCACCAATTTACCTCCATCTTTGSCACATTT 720  
Qy 721 TGGGCCCCCGCATCATCATGATCTTTTACCACCTCTATGGGGGCGCCCATCCAGAACCGC 780  
Db 721 TGGGCCCCCGCATCATCATGATCTTTTACCACCTCTATGGGGGCGCCCATCCAGAACCGC 780  
Qy 781 TGGCTGGTACACATCATGTCGACATTCGCCAACATGCTAGCCCTTCTGAAACAGCCATC 840  
Db 781 TGGCTGGTACACATCATGTCGACATTCGCCAACATGCTAGCCCTTCTGAAACAGCCATC 840  
Qy 841 AACTTCTTCTCTACTGCTTCTACAGCAAGCGGTTCCGACCATGGCAGCGCCAGCTC 900  
Db 841 AACTTCTTCTCTACTGCTTCTACAGCAAGCGGTTCCGACCATGGCAGCGCCAGCTC 900  
Qy 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC 960

```
Db 901 AAGGCTTTCTTCAAGTGGCAGAAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960
Qy 961 ATAACAAGTAGCCCTGTGATCTCGCGGCAAACTCACAAGTGCATCAAGATGCTGTGTAC 1020
Db 961 ATAACAAGTAGCCCTGTGATCTCGCGGCAAACTCACAAGTGCATCAAGATGCTGTGTAC 1020
Qy 1021 CAGTATGACAAAATGGAAAACCTATAAAA 1050
Db 1021 CAGTATGACAAAATGGAAAACCTATAAAA 1050

RESULT 11
ACCT71785
ID ACCT71785 standard; DNA; 1062 BP.
XX AC ACCT71785;
XX DT 25-JUL-2003 (first entry)
XX Human G protein-coupled receptor DNA #SEQ ID 1.
XX Human; G protein-coupled receptor; anorectic; eating disorder; obesity;
KW gene; ds.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT 1..1062
FT /*tag= a
FT /product= "G protein-coupled receptor"
XX WO2003027142-A1.
XX PD 03-APR-2003.
XX PF 19-SEP-2002; 2002WO-JP009626.
XX PR 21-SEP-2001; 2001JP-00288278.
XX PA (YAMA ) YAMANOUCHI PHARM CO LTD.
XX PI Matsumoto S, Takasaki J, Kurama T, Saito T, Kamohara M, Soga T;
XX PI Hiyaama H;
XX PI WPI; 2003-333291/31.
XX DR P-PSDB; ABR44437.
XX PT New G protein-coupled receptor and encoded gene, useful in screening
XX PT preventives or remedies for eating disorders or obesity.
XX PS Disclosure; Page 31-33; 48pp; Japanese.
XX CC The invention relates to a novel G protein-coupled receptor. The protein
XX CC and its encoded gene are useful for screening preventives or remedies for
XX CC eating disorders or obesity. The current sequence represents a G protein
XX CC coupled receptor encoding sequence
XX SQ Sequence 1062 BP; 241 A; 342 C; 207 G; 272 T; 0 U; 0 Other;

Query Match 93.8%; Score 1050; DB 10; Length 1062;
Best Local Similarity 100.0%; Pred. No. 4.9e-287;
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGAGCACAGCAGCCGACCTCGCAGCAACAGCTCGTGTCTTGTGGTTCCTCCCGGC 60
Db 1 ATGAGAGCACAGCAGCCGACCTCGCAGCAACAGCTCGTGTCTTGTGGTTCCTCCCGGC 60
Qy 61 TCGGCTCGGCTTGGGTTTGGTCCCGTGTCTACTACAGCTCTTCTGCTGTGCTCGGT 120
Db 61 TCGGCTCGGCTTGGGTTTGGTCCCGTGTCTACTACAGCTCTTCTGCTGTGCTCGGT 120
Qy 121 TTACCAGCAAAATATCTTACAGTGTATCATCTCTCCAGCTGGTGGCAAGAAGACAAG 180
Db 121 TTACCAGCAAAATATCTTACAGTGTATCATCTCTCCAGCTGGTGGCAAGAAGACAAG 180
```

```
Db 121 TTACCAGCAAAATATCTTACAGTGTATCATCTCTCCAGCTGGTGGCAAGAAGACAAG 180
Qy 181 TCCTCTTACAACTATCTTGGCACTGGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 240
Db 181 TCCTCTTACAACTATCTTGGCACTGGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 240
Qy 241 GTGTTTGTGGACTTCTTGGTGGAAAGATTTCATCTTGAACATGACAGATGCCTCAGGTCCCC 300
Db 241 GTGTTTGTGGACTTCTTGGTGGAAAGATTTCATCTTGAACATGACAGATGCCTCAGGTCCCC 300
Qy 301 GACAAGATCATAGAAGTCTGGAAATCTCATCCATCCACACCTCCATATGATTACTGTA 360
Db 301 GACAAGATCATAGAAGTCTGGAAATCTCATCCATCCACACCTCCATATGATTACTGTA 360
Qy 361 CCGTTAACCAATTGACAGGTATATCGCTGTCTGCGCACCGGCTCAAGTACCACACCGTCTCA 420
Db 361 CCGTTAACCAATTGACAGGTATATCGCTGTCTGCGCACCGGCTCAAGTACCACACCGTCTCA 420
Qy 421 TACCAGCCGCGCACCGGAAAGTCAATTGTAAGTGTTTTACATCAGCTGCTTCTGACCAAGC 480
Db 421 TACCAGCCGCGCACCGGAAAGTCAATTGTAAGTGTTTTACATCAGCTGCTTCTGACCAAGC 480
Qy 481 ATCCCTTATTAATCTGCTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540
Db 481 ATCCCTTATTAATCTGCTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540
Qy 541 CAGTCTCATCTGGATCCACTGCTTCAACGCTCTACCTGGTGGCTGCTGCTCTCTTCTTC 600
Db 541 CAGTCTCATCTGGATCCACTGCTTCAACGCTCTACCTGGTGGCTGCTGCTCTCTTCTTC 600
Qy 601 ATCTTGAACATCAATCTTGTGTACAGCTCAGGAGGAAGAGCAATTTTCGCTCTCGTGGC 660
Db 601 ATCTTGAACATCAATCTTGTGTACAGCTCAGGAGGAAGAGCAATTTTCGCTCTCGTGGC 660
Qy 661 TACTCCAGGGGGAAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720
Db 661 TACTCCAGGGGGAAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720
Qy 721 TGGGCCCCCGGCATCATCATGATTTTACACCTCTATGGGGGCGCCCATCCAGAACCGC 780
Db 721 TGGGCCCCCGGCATCATCATGATTTTACACCTCTATGGGGGCGCCCATCCAGAACCGC 780
Qy 781 TGGCTGGTACACATCATGTTCGACATTTGCCAAACATGCTAGCCCTTCTGAAACACAGCCATC 840
Db 781 TGGCTGGTACACATCATGTTCGACATTTGCCAAACATGCTAGCCCTTCTGAAACACAGCCATC 840
Qy 841 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTCCGCAACCATGGCAGCGCCCAAGCTC 900
Db 841 AACTTCTTCTTACTGCTTTCATCAGCAAGCGGTTCCGCAACCATGGCAGCGCCCAAGCTC 900
Qy 901 AAGGCTTCTTCAAGTGGCAGAAGCAACCTGTACAGTTCTACACCAATCATACTTTTCC 960
Db 901 AAGGCTTCTTCAAGTGGCAGAAGCAACCTGTACAGTTCTACACCAATCATACTTTTCC 960
Qy 961 ATAAACAAGTAGCCCTGTGATCTCGCGGCAAACTCACAAGTGCATCAAGATGCTGTGTAC 1020
Db 961 ATAAACAAGTAGCCCTGTGATCTCGCGGCAAACTCACAAGTGCATCAAGATGCTGTGTAC 1020
Qy 1021 CAGTATGACAAAATGGAAAACCTATAAAA 1050
Db 1021 CAGTATGACAAAATGGAAAACCTATAAAA 1050

RESULT 12
ABV73364
ID ABV73364 standard; DNA; 1202 BP.
XX AC ABV73364;
XX XX 22-JAN-2003 (first entry)
XX DE Human TGR20 polypeptide encoding DNA.
XX XX
```

KW G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory;  
 KW antilipemic; neurotropic; neuroprotective; antianemic; antiulcer; human;  
 XX antiparkinsonian; antilipemic; TGR20; gene; ds.  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers  
 FT CDS 68..1129  
 FT /\*tag= a  
 FT /product= "TGR20"  
 FT /note= "GPCR polypeptide"  
 XX  
 PN WO200277001-A2.  
 XX  
 XX 03-OCT-2002.  
 XX  
 XX 08-MAR-2002; 2002WO-US007171.  
 XX  
 XX 09-MAR-2001; 2001US-00802803.  
 PR 16-MAR-2001; 2001US-0276649P.  
 XX  
 XX (TULA-) TULARIK INC.  
 XX  
 XX Tian H, Zhao J, Chen J, Cutler G;  
 XX  
 DR WPI; 2003-018881/01.  
 DR P-PSDB; ABB82499.  
 XX  
 XX New G-protein coupled receptor polypeptides and polynucleotides useful  
 PT for identifying compounds for treating a TGR-associated disorder, e.g.  
 PT psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's  
 PT disease, anemia.  
 XX  
 XX Claim 6; Page 63; 87pp; English.  
 XX  
 XX The invention relates to G-protein coupled receptor (GPCR) polypeptides  
 CC and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183,  
 CC TGR341, TGR211, and TGR79. The polypeptides and nucleic acids are  
 CC useful for identifying compounds for treating a TGR-associated disorder,  
 CC such as psoriasis, inflammatory bowel disease, hyperlipidemia,  
 CC Parkinson's disease, Huntington's disease, anemia, immune and blood  
 CC disorders, ulcerative colitis, Crohn's disease or spleen enlargement.  
 CC They are also useful for identifying cells such as kidney, liver,  
 CC hypothalamus, colon, adipose, or spleen cells, for forensics and,  
 CC paternity determination, diagnosing diseases and examining signal  
 CC transduction. The present sequence represents a human TGR20 polypeptide  
 CC encoding DNA  
 XX  
 XX Sequence 1202 BP; 267 A; 380 C; 258 G; 297 T; 0 U; 0 Other;  
 SQ  
 Query Match 93.8%; Score 1050; DB 8; Length 1202;  
 Best Local Similarity 100.0%; Pred. No. 5,2e-287;  
 Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ATGGAGCACAGCAGCCCACTGGCAGCAACAGCTCGCTGCTTGTGGTCCCGCGC 60  
 DB 68 ATGGAGCACAGCAGCCCACTGGCAGCAACAGCTCGCTGCTTGTGGTCCCGCGC 127  
 QY 61 TCGGCTCGCGCTTGGGTTTGGTGGCGGTGCTTACTACAGCTCTTGTGTGCTCGGT 120  
 DB 128 TCGGCTCGCGCTTGGGTTTGGTGGCGGTGCTTACTACAGCTCTTGTGTGCTCGGT 187  
 QY 121 TTACACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 180  
 DB 188 TTACACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 247  
 QY 181 TCTCTCTACAACTATCTTTGGCAGCTCGTGTGCGGACATCTTGTCTCTTTTCATA 240  
 DB 248 TCTCTCTACAACTATCTTTGGCAGCTCGTGTGCGGACATCTTGTCTCTTTTCATA 307  
 QY 241 GTGTTTGTGAGCTTCTCTGTTGGAAGATTTTCATCTTGAACTGAGTGCCTCAGGTCCC 300  
 DB 308 GTGTTTGTGAGCTTCTCTGTTGGAAGATTTTCATCTTGAACTGAGTGCCTCAGGTCCC 367

QY 301 GACAAGATCATAGAAAGTCTGGAATTCTCATCCATCCACACCTCCATATGGAATTACTGTA 360  
 DB 368 GACAAGATCATAGAAAGTCTGGAATTCTCATCCATCCACACCTCCATATGGAATTACTGTA 427  
 QY 361 CCGTTAACCATTTGACAGGTATATCGCTGTGCGACCCGCTCAAGTACCACAGGCTCTCA 420  
 DB 428 CCGTTAACCATTTGACAGGTATATCGCTGTGCGACCCGCTCAAGTACCACAGGCTCTCA 487  
 QY 421 TACCAGCCCGCACCCGGAAAGTCAATTGTAAGTGTTTTACATCACCTGCTTCTTGACACGAC 480  
 DB 488 TACCAGCCCGCACCCGGAAAGTCAATTGTAAGTGTTTTACATCACCTGCTTCTTGACACGAC 547  
 QY 481 ATCCCTTATTACTGCTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
 DB 548 ATCCCTTATTACTGCTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 607  
 QY 541 CAGCTCCTCATCTGATCCATGCTTCAACGCTTACCTGCTGCTGCTGCTGCTGCTGCTTTC 600  
 DB 608 CAGCTCCTCATCTGATCCATGCTTCAACGCTTACCTGCTGCTGCTGCTGCTGCTTTC 667  
 QY 601 ATCTTGAATCAATCATTTGTGTACAAAGCTCAGAGAGAGAGCAATTTTCTGCTCCGCTGC 660  
 DB 668 ATCTTGAATCAATCATTTGTGTGTACAAAGCTCAGAGAGAGAGCAATTTTCTGCTCCGCTGC 727  
 QY 661 TACTCCACGGGGAAGACCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720  
 DB 728 TACTCCACGGGGAAGACCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 787  
 QY 721 TGGGCCCCCGGATCATCATGATTTTACACCTCTATGCGGCGCCCATCCAGAACCCG 780  
 DB 788 TGGGCCCCCGGATCATCATGATTTTACCACTCTATGCGGCGCCCATCCAGAACCCG 847  
 QY 781 TGGCTGGTACACATCATGTGCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCATC 840  
 DB 848 TGGCTGGTACACATCATGTGCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCATC 907  
 QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATCGCAGCCCGCAGCTC 900  
 DB 908 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATCGCAGCCCGCAGCTC 967  
 QY 901 AAGGCTTCTTCAAGTGGCAGAGCAACCTGTACAGTTCTACACCAATCAATCAATCTTTC 960  
 DB 968 AAGGCTTCTTCAAGTGGCAGAGCAACCTGTACAGTTCTACACCAATCAATCAATCTTTC 1027  
 QY 961 ATAACAAGTAGCCCTGATCTCGCCGCAAACTCAGACTGATCAAGATGCTGGTGTAC 1020  
 DB 1028 ATAACAAGTAGCCCTGATCTCGCCGCAAACTCAGACTGATCAAGATGCTGGTGTAC 1087  
 QY 1021 CAGTATGACAAAAATGAAAACTATAAAA 1050  
 DB 1088 CAGTATGACAAAAATGAAAACTATAAAA 1117  
 RESULT 13  
 ABK15563  
 ID ABK15563 standard; cDNA; 1062 BP.  
 XX  
 AC ABK15563;  
 XX  
 DT 08-MAY-2002 (first entry)  
 XX  
 DE cDNA encoding novel G-protein coupled receptor TGR8, version #2.  
 KW G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;  
 KW cystotic; antiinflammatory; antiulcer; fetal brain;  
 KW central nervous system disease; circulatory organ disorder; cancer;  
 KW metabolic disease; immunological disease; gastrointestinal disease;  
 KW gene therapy; transgenic animal; human; gene; ss.  
 OS Homo sapiens.  
 XX  
 PH Key Location/Qualifiers



FT CDS 1. .1062  
FT /\*tag= a  
FT /product= "TGR8"  
FT /note= "G-protein coupled receptor"  
XX  
PN WO200194582-A1.  
PD 13-DEC-2001.  
XX  
XX 01-JUN-2001; 2001WO-JP004643.  
XX 02-JUN-2000; 2000JP-00170446.  
PR 23-JUN-2000; 2000JP-00194926.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Terao Y, Matsui H, Shintani Y;  
PI WPI; 2002-164317/21.  
DR P-PSDB; AAU76416.  
XX  
PT Human fetal brain-originated G protein-coupled receptor protein TGR8 and  
PT encoding DNA, for developing drugs to treat e.g. diseases of the central  
PT nervous system or circulatory organs, cancer, and metabolic diseases.  
XX  
PS Claim 5; Page 94-95; 102pp; Japanese.  
XX  
XX The invention describes a human fetal brain-originated G protein-coupled  
CC receptor protein, or its salt. The protein and encoded DNA are useful for  
CC developing drugs to treat e.g. diseases of the central nervous system or  
CC circulatory organs, cancer, metabolic diseases, immunological diseases  
CC and gastrointestinal diseases. The invention also describes creation of a  
CC probe for gene therapy and construction of a transgenic animal. This  
CC sequence encodes the novel G-protein coupled receptor TGR8, described in  
CC the method of the invention  
XX  
SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;  
  
Query Match 93.7%; Score 1048.4; DB 6; Length 1062;  
Best Local Similarity 99.9%; Pred. No. 1.4e-286;  
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 ATGAGCAGCAGCAGCCACCTCGCAGCCAAACAGCTCGCTGCTTTGGTGTCCCGCGC 60  
Db 1 ATGAGCAGCAGCAGCCACCTCGCAGCCAAACAGCTCGCTGCTTTGGTGTCCCGCGC 60  
  
QY 61 TCGGCTCGGGCTTGGGTTTGGTGGCCGTGTTACTACAGCCTTCTGCTGTGCTCGGT 120  
Db 61 TCGGCTCGGGCTTGGGTTTGGTGGCCGTGTTACTACAGCCTTCTGCTGTGCTCGGT 120  
  
QY 121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 180  
Db 121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 180  
  
QY 181 TCCTCCTACAACTATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTCTTTTCATA 240  
Db 181 TCCTCCTACAACTATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTCTTTTCATA 240  
  
QY 241 GTGTTTGTGGACTTCCTGTGGAAGATTTCATTTTGAACATGCAAGTCCCTCAGGTCCCC 300  
Db 241 GTGTTTGTGGACTTCCTGTGGAAGATTTCATTTTGAACATGCAAGTCCCTCAGGTCCCC 300  
  
QY 301 GACAGATCATAGAGTCTGGAAATTCATCCATCCACACCTCCATATGATGATCTGTA 360  
Db 301 GACAGATCATAGAGTCTGGAAATTCATCCATCCACACCTCCATATGATGATCTGTA 360  
  
QY 361 CGGTTAACATTGACAGGTATATCGCTGTGCGCACCGCTCAAGTACACACGCTCTCA 420  
Db 361 CGGTTAACATTGACAGGTATATCGCTGTGCGCACCGCTCAAGTACACACGCTCTCA 420  
  
QY 421 TACCAGCCCGCACC CGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480  
Db 421 TACCAGCCCGCACC CGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480

QY 481 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
Db 481 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
  
QY 541 CACGTCTCATCTGGATCCACTGCTTCAACGCTGTACCTGGTGGCCCTGTCTCCATCTTCTTC 600  
Db 541 CACGTCTCATCTGGATCCACTGCTTCAACGCTGTACCTGGTGGCCCTGTCTCCATCTTCTTC 600  
  
QY 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTCTGTCCTGGGC 660  
Db 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTCTGTCCTGGGC 660  
  
QY 661 TACTCCACGGGGAAGACCGCCATCTTGTTCACCATTAACCTCCATCTTTGCCACACTT 720  
Db 661 TACTCCACGGGGAAGACCGCCATCTTGTTCACCATTAACCTCCATCTTTGCCACACTT 720  
  
QY 721 TGGGCCCCCGGCATCATCATGATTTTACACCTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 721 TGGGCCCCCGGCATCATCATGATTTTACACCTCTATGGGGCGCCCATCCAGAACCGC 780  
  
QY 781 TGGTGTGTACACATCATGTTCGCAATTGCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
Db 781 TGGTGTGTACACATCATGTTCGCAATTGCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
  
QY 841 AACTTTCTTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACCTATGGCAGCGCCACGCTC 900  
Db 841 AACTTTCTTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACCTATGGCAGCGCCACGCTC 900  
  
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960  
  
QY 961 ATAAACAAGTAGCCCTCGATCTCGCGGCAAACTCACTGCAATCAAGATGCTGGTGTAC 1020  
Db 961 ATAAACAAGTAGCCCTCGATCTCGCGGCAAACTCACTGCAATCAAGATGCTGGTGTAC 1020  
  
QY 1021 CAGTATGCAAAAATGGAACCTATATAA 1050  
Db 1021 CAGTATGCAAAAATGGAACCTATATAA 1050  
  
RESULT 14  
ABT04873  
ID ABT04873 standard; cDNA; 1062 BP.  
XX  
AC ABT04873;  
XX  
DT 11-OCT-2002 (first entry)  
XX  
DE Human G protein coupled receptor hrUP35 coding sequence.  
XX  
XX Human; G-protein coupled receptor; GPCR; hrUP28; hrUP29; hrUP30; hrUP31;  
KW hrUP32; hrUP33; hrUP34; hrUP35; hrUP36; hrUP37; gene; ss.  
XX Homo sapiens.  
XX WO200242461-A2.  
FN 30-MAY-2002.  
XX  
XX 26-NOV-2001; 2001WO-US044386.  
XX  
PR 27-NOV-2000; 2000US-0253404P.  
PR 12-DEC-2000; 2000US-0255366P.  
PR 20-FEB-2001; 2001US-0270266P.  
PR 20-FEB-2001; 2001US-0270286P.  
PR 06-APR-2001; 2001US-0282032P.  
PR 06-APR-2001; 2001US-0282356P.  
PR 06-APR-2001; 2001US-0282358P.  
PR 14-MAY-2001; 2001US-0290917P.  
PR 31-JUL-2001; 2001US-0309208P.



XX (AREN-) ARENA PHARM INC.  
 XX Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;  
 XX WPI; 2002-566565/60.  
 XX P-PSDB; ABJ04075.  
 XX Novel endogenous and non-endogenous versions of G protein-coupled  
 PT receptor useful for identification of candidate compounds as receptor  
 PT agonists or antagonists for use as therapeutic agents.  
 XX Claim 31; Page 70; 84pp; English.  
 XX The present invention provides the protein and coding sequences of  
 CC several human G-protein coupled receptors (GPCRs). These can be used in  
 CC the identification of candidate compounds as receptor agonists or inverse  
 CC agonists having applicability as therapeutic agents. The present sequence  
 CC is a GPCR coding sequence of the invention  
 XX  
 SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;  
 Query Match 93.7%; Score 1048.4; DB 6; Length 1062;  
 Best Local Similarity 99.9%; Pred. No. 1.4e-286;  
 Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGGAGCACACGACGCCACCTCGCAGCAACAGCTCGCTGCTTGTGGTGGTCCCGCGC 60  
 DB 1 ATGGAGCACACGACGCCACCTCGCAGCAACAGCTCGCTGCTTGTGGTGGTCCCGCGC 60  
 QY 61 TGGGCTGGGCTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120  
 DB 61 TGGGCTGGGCTGGGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120  
 QY 121 TTACGAGCAATATCTTGACAGTATCTCCAGCTCCAGCTGGTGGTGGTGGTGGTGGTGGT 180  
 DB 121 TTACGAGCAATATCTTGACAGTATCTCCAGCTCCAGCTGGTGGTGGTGGTGGTGGTGGT 180  
 QY 181 TCTCTTACAACTATCTCTTGGCAGCTCGCTGCTGGCAGACATCTTGGTCTCTTTTTCATA 240  
 DB 181 TCTCTTACAACTATCTCTTGGCAGCTCGCTGCTGGCAGACATCTTGGTCTCTTTTTCATA 240  
 QY 241 GTGTTTGGGACTTCTGTGTGAAGATTTCACTTGAACATGACAGATGCTCAGGTGCC 300  
 DB 241 GTGTTTGGGACTTCTGTGTGAAGATTTCACTTGAACATGACAGATGCTCAGGTGCC 300  
 QY 301 GACAGATCATAGAAGTCTGGATTTCTATCCATCCACCTCCATATGATGATCTGTA 360  
 DB 301 GACAGATCATAGAAGTCTGGATTTCTATCCATCCACCTCCATATGATGATCTGTA 360  
 QY 361 CGTTAAACCATTCACAGGTATATCGTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 420  
 DB 361 CGTTAAACCATTCACAGGTATATCGTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 420  
 QY 421 TACCCAGCCCGCACCCGGAAGTCAATTGAAGTGTATATCACTGCTTCTGACCCAGC 480  
 DB 421 TACCCAGCCCGCACCCGGAAGTCAATTGAAGTGTATATCACTGCTTCTGACCCAGC 480  
 QY 481 ATCCCTTATATCTGTGGGCGCAACATCTGCACTGAAGACTATCAGACCTCTGTGCAT 540  
 DB 481 ATCCCTTATATCTGTGGGCGCAACATCTGCACTGAAGACTATCAGACCTCTGTGCAT 540  
 QY 541 CAGTCTCTATCTGGATCACTGCTTCCAGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 DB 541 CAGTCTCTATCTGGATCACTGCTTCCAGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
 QY 601 ATCTTGAATCAATCTGTGTGAAGTCAAGTCAAGGAGGAGCAATTTTCTGCTCCGTGGC 660  
 DB 601 ATCTTGAATCAATCTGTGTGAAGTCAAGTCAAGGAGGAGCAATTTTCTGCTCCGTGGC 660  
 QY 661 TACTCCACGGGGAAGACCAACCGGCACTTTGTTTCCATTTACCTTCCATCTTTGCCACACT 720  
 DB 661 TACTCCACGGGGAAGACCAACCGGCACTTTGTTTCCATTTACCTTCCATCTTTGCCACACT 720

QY 721 TGGGCCCCCGCATCATCATGATTTTACACCTTATGGGGGCGCCCATCCAGAACCGC 780  
 DB 721 TGGGCCCCCGCATCATCATGATTTTACACCTTATGGGGGCGCCCATCCAGAACCGC 780  
 QY 781 TGGCTGGGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAAACACAGCCATC 840  
 DB 781 TGGCTGGGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAAACACAGCCATC 840  
 QY 841 AACTTTCTTCTTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGCAGCCGCTC 900  
 DB 841 AACTTTCTTCTTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGCAGCCGCTC 900  
 QY 901 AAGGCTTTCTTCAAGTGGCAGAACCTGTACAGTTCTACCAATCATTAACCTTTCC 960  
 DB 901 AAGGCTTTCTTCAAGTGGCAGAACCTGTACAGTTCTACCAATCATTAACCTTTCC 960  
 QY 961 ATAACAAGTAGCCCTGATCTCGCCGCAAACTCAGACTGATCAAGATGCTGGTGTAC 1020  
 DB 961 ATAACAAGTAGCCCTGATCTCGCCGCAAACTCAGACTGATCAAGATGCTGGTGTAC 1020  
 QY 1021 CAGTATGACAAAATGGAACCTATAAAA 1050  
 DB 1021 CAGTATGACAAAATGGAACCTATAAAA 1050  
 RESULT 15  
 ACC44115  
 ID ACC44115 standard; DNA; 1062 BP.  
 XX ACC44115;  
 XX 27-JUN-2003 (first entry)  
 XX Human AXOR-57 coding sequence.  
 DE ds; antibacterial; antifungal; antiviral; antiprotazoal; analgesic; gene;  
 KW cytosolic; antidiabetic; anorectic; anabolic; antiaesthetic; antagonist;  
 KW antiparkinsonian; cardiac; hypotensive; hypertensive; nephrotropic;  
 KW osteopathic; antianal; antitumor; antiallergic; antimigraine; agonist;  
 KW antileptic; norepinephrine; tranquilizer; neuroprotective; antidepressant;  
 KW vaccine; gene therapy; G-protein coupled receptor; receptor; infection;  
 KW pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;  
 KW Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;  
 KW vomiting; schizophrenia; depression; dementia; Huntington's disease.  
 XX Homo sapiens.  
 OS  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..1062  
 FT /\*tag= a  
 FT /product= "AXOR-57 protein"  
 XX  
 PN GB2365009-A.  
 XX  
 PD 13-FEB-2002.  
 XX  
 PF 11-APR-2001; 2001GB-00009018.  
 XX  
 PR 11-APR-2001; 2001GB-00009018.  
 XX  
 PA (SMIK ) SMITHKLINE BEECHAM CORP.  
 PA (SMIK ) SMITHKLINE BEECHAM PLC.  
 XX  
 PI Elshourbagy N, Gattu M, Michalovich D, Shabon U;  
 DR WPI; 2003-203569/20.  
 DR P-PSDB; ABP98724.  
 XX  
 PT New G-protein coupled receptor, AXOR 57, for diagnosing and treating  
 PT diseases, such as, pain, cancer, diabetes, obesity, anorexia, asthma,  
 PT Parkinson's disease, hypotension, hypertension, urinary retention, and  
 PT osteoporosis.

XX  
PS Claim 2; Page 27; 32pp; English.

The invention relates to the isolation of the coding sequence of a novel G-protein coupled (7TM) receptor AXOR 95 (this sequence) or fragment encoding a polypeptide having at least 95 % identity with AXOR-57 protein. The protein has been shown to have homology to the human chromosome 16 clone C17B-H1.036A2 (AC008785). The protein and polynucleotides encoding the protein are used in screening for compounds that stimulate or inhibit the function or level of the polypeptide or polynucleotides, such as, agonists and antagonists. The protein, polynucleotides, and antibodies to the protein are used in diagnostic kits, to diagnose a disease. Polynucleotides encoding the protein are used for chromosome localization studies, or for tissue expression studies. The protein and nucleic acids encoding the protein are used in vaccines for treating diseases such as, bacterial, fungal, protozoal, and viral infections, pain, cancer, diabetes, obesity, anorexia, bulimia, asthma, Parkinson's disease, acute heart failure, hypotension, hypertension, urinary retention, osteoporosis, angina pectoris, myocardial infarction, stroke, ulcer, allergy, benign prostatic hypertrophy, migraine, vomiting, anxiety, schizophrenia, depression, delirium, dementia, severe mental retardation, or Huntington's disease

SQ Sequence 1062 BP; 240 A; 342 C; 208 G; 272 T; 0 U; 0 Other;

```
Query Match      93.7%; Score 1048.4; DB 10; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.4e-286;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

Qy	1	ATGGAGCACACGACGCGCCACTCGCAGCCCAACAGCTCGCTGCTTTGGTGGTCCCCCGCG	60
Db	1	ATGGAGCACACGACGCGCCACTCGCAGCCCAACAGCTCGCTTTGGTGGTCCCCCGCG	60
Qy	61	TGGGCTCGGGCTTGGGTTTCGTGCCCGTGTCTACTACAGCCTCTTGTGTGCTTCGGT	120
Db	61	TGGGCTCGGGCTTGGGTTTCGTGCCCGTGTCTACTACAGCCTCTTGTGTGCTTCGGT	120
Qy	121	TTACAGAGCAAAATATTCTTGACAGTGANATCTCTCCAGCTGTGTGGCAGAAGACAGAAG	180
Db	121	TTACAGAGCAAAATATTCTTGACAGTGANATCTCTCCAGCTGTGTGGCAGAAGACAGAAG	180
Qy	181	TCTCTCTCAAACTATCTCTTGCGCACTGCTGCTGCGACATCTTGGTCTCTTTTTCATA	240
Db	181	TCTCTCTCAAACTATCTCTTGCGCACTGCTGCTGCGACATCTTGGTCTCTTTTTCATA	240
Qy	241	GTGTTTGGGACTTCCTGTGGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCGCC	300
Db	241	GTGTTTGGGACTTCCTGTGGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCGCC	300
Qy	301	GACAGATCATAGAAGTCTGGAAATTCATCTCATCTCCATCACACCTCCATATGGATTACTGT	360
Db	301	GACAGATCATAGAAGTCTGGAAATTCATCTCATCTCCATCACACCTCCATATGGATTACTGT	360
Qy	361	CGGTTAACCATTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACACACGGTCTCA	420
Db	361	CGGTTAACCATTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACACACGGTCTCA	420
Qy	421	TACCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTTTACATCACCTGCTTCTTGACACG	480
Db	421	TACCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTTTACATCACCTGCTTCTTGACACG	480
Qy	481	ATCCCCGTATTACTGTGTGCGCCCAACATCTGGAGCTGAAGACTACATCAGCACCTCTGTGCAT	540
Db	481	ATCCCCGTATTACTGTGTGCGCCCAACATCTGGAGCTGAAGACTACATCAGCACCTCTGTGCAT	540
Qy	541	CAGGTCTCATCTGGATCCACTGTCTTACCCTGTCTACCTGTGTGCCCTGTCTCACTTCTTTC	600
Db	541	CAGGTCTCATCTGGATCCACTGTCTTACCCTGTCTACCTGTGTGCCCTGTCTCACTTCTTTC	600
Qy	601	ATCTTGAATCTCAATCATTTGTGTACAAGCTCAGGAGGAAGACAATTTTGTCTTCGTGGC	660
Db	601	ATCTTGAATCTCAATCATTTGTGTACAAGCTCAGGAGGAAGACAATTTTGTCTTCGTGGC	660

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 08:34:56 ; Search time 4812 Seconds  
(without alignments)  
11267.952 Million cell updates/sec

Title: US-10-689-832-19  
Perfect score: 1119  
Sequence: 1 atggagcacacgacgccc.....cttggtcatcatctgtga 1119

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl.\*

- 1: gb\_ba.\*
- 2: gb\_hlg.\*
- 3: gb\_in.\*
- 4: gb\_om.\*
- 5: gb\_ov.\*
- 6: gb\_pat.\*
- 7: gb\_ph.\*
- 8: gb\_pl.\*
- 9: gb\_pr.\*
- 10: gb\_ro.\*
- 11: gb\_sts.\*
- 12: gb\_sy.\*
- 13: gb\_un.\*
- 14: gb\_vi.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	100.0	1119	6	AX254975 Sequence
2	1119	100.0	1130	6	AX375235 Sequence
3	1050	93.8	1059	6	AX463227 Sequence
4	1050	93.8	1062	6	BD183203 Novel G p
5	1050	93.8	1062	6	BD103851 Novel G p
6	1049.6	93.8	1167	9	AY635179 Homo sapi
7	1048.4	93.7	1062	6	BD183204 Novel G p
8	1048.4	93.7	1062	6	AX498192 Sequence
9	1048.4	93.7	1062	6	BD103852 Novel G p
10	1048.4	93.7	1343	6	AX254977 Sequence
11	1048.4	93.7	2117	6	CQ867782 Sequence
12	1040.4	93.0	1062	6	AX664703 Sequence
13	1040.4	93.0	1526	6	AX664701 Sequence
14	934.8	83.5	1062	6	AX463235 Sequence
15	929.8	83.1	957	6	CQ737667 Sequence
16	926	82.8	1002	6	AX375442 Sequence
17	926	82.8	1158	6	BD144286 Novel G-p
18	926	82.8	1167	6	CQ848070 Sequence
19	926	82.8	1167	9	AB083594 Homo sapi

C	20	926	82.8	101882	9	AC021089	AC021089 Homo sapi
	21	924.8	82.6	1000	6	AX342465	AX342465 Sequence
	22	924.4	82.6	1788	6	CQ739127	CQ739127 Sequence
C	23	924.4	82.6	110000	2	AL139235_0	AL139235 Homo sapi
	24	922.4	82.4	1032	6	AX463231	AX463231 Sequence
	25	922.4	82.4	1070	6	AX463230	AX463230 Sequence
	26	922.4	82.4	1826	6	AX463232	AX463232 Sequence
C	27	844.2	75.4	864	6	AX230119	AX230119 Sequence
	28	811.2	72.5	963	6	AX657444	AX657444 Sequence
	29	776.4	69.4	183326	10	AC124438	AC124438 Mus muscu
C	30	767.6	68.6	140885	2	AC119798	AC119798 Rattus no
	31	767.6	68.6	203372	2	AC121346	AC121346 Rattus no
C	32	767.6	68.6	230813	2	AC114097	AC114097 Rattus no
	33	756.4	67.6	795	6	AX549301	AX549301 Sequence
	34	756.4	67.6	795	9	HS4303165	AJ303165 Homo sapi
	35	548	49.0	558	9	AY255545	AY255545 Homo sapi
C	36	524	46.8	930	6	AX230116	AX230116 Sequence
	37	443.6	39.6	1149	5	AY288414	AY288414 Takifugu
	38	406.4	36.3	242526	2	CR846080	CR846080 Danio rer
	39	268.2	24.0	945	9	AY255622	AY255622 Homo sapi
	40	268.2	24.0	1962	6	AX375239	AX375239 Sequence
	41	266.6	23.8	1122	6	BD140885	BD140885 Novel G p
	42	266.6	23.8	1122	6	BD185294	BD185294 Novel G p
	43	266.6	23.8	1125	6	AX498182	AX498182 Sequence
	44	266.6	23.8	1125	6	AX664706	AX664706 Sequence
	45	266.6	23.8	1389	9	AY288421	AY288421 Homo sapi

ALIGNMENTS

RESULT 1  
AX254975  
LOCUS AX254975 1119 bp DNA linear PAT 10-OCT-2001  
DEFINITION Sequence 19 from Patent WO0170978.  
ACCESSION AX254975  
VERSION AX254975.1 GI:16074503  
KEYWORDS Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1  
AUTHORS Taupier,R.J., Majumder,K., Spaderna,S.K., Smithson,G., Mezes,P.S. and Vernet,C.A.  
TITLE Polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0170978-A 19 27-SEP-2001;  
Curagen Corporation (US)  
FEATURES  
source  
1. .1119  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN

Query Match 100.0%; Score 1119; DB 6; Length 1119;  
Best Local Similarity 100.0%; Pred. No. 3.4e-269;  
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACACGACGCCACCTCGCAGCCAAACAGCTCGCTGCTTGGTGGTCCCCGGC 60  
DB 1 ATGGAGCACACGACGCCACCTCGCAGCCAAACAGCTCGCTGCTTGGTGGTCCCCGGC 60  
QY 61 TCGGCTGGGGTTGGTGGTTCGTCGCCGGTGTCTACTACAGCCTCTTGTGTGCTTCGGT 120  
DB 61 TCGGCTGGGGTTGGTGGTTCGTCGCCGGTGTCTACTACAGCCTCTTGTGTGCTTCGGT 120  
QY 121 TTACACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
DB 121 TTACACGACAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
QY 181 TCGTCTTACAACTATCTCTTGGCACTCGCTGCTGGGACATCTTGTCTCTTTTCATA 240

181	Db	TCCTCTCAAACTATCTCTTGGCACTCGTGCTGCTCGCGACATCTTTGGTCCCTCTTTTTTCATA	240
241	Qy	GTGTTTGTGGACTTCCTGTTTGGAAAGATTTTCATCTTTGAACATGCAGATGCCTCAGGTCCCC	300
241	Db		
241	Db	GTGTTTGTGGACTTCCTGTTTGGAAAGATTTTCATCTTTGAACATGCAGATGCCTCAGGTCCCC	300
301	Qy	GACAAGATCATAGAAGTGTCTGGAATTTCTCATPCATCCATCCACACCTCCATATGGAATTCGTGA	360
301	Db	GACAAGATCATAGAAGTGTCTGGAATTTCTCATPCATCCATCCACACCTCCATATGGAATTCGTGA	360
361	Qy	CCGTTAACCATTTGACAGGTATATCGCTGTCTGCCACACCGGTCAAGTACCAACGGTCTCA	420
361	Db	CCGTTAACCATTTGACAGGTATATCGCTGTCTGCCACACCGGTCAAGTACCAACGGTCTCA	420
421	Qy	TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTTTACATCACTGCTCTCTCCAGCAGC	480
421	Db	TACCCAGCCGACCCGGAAGTCAATTTGAAGTGTTTACATCACTGCTCTCTCCAGCAGC	480
481	Qy	ATCCCCATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT	540
481	Db	ATCCCCATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT	540
541	Qy	CAGTCTCATCTGGATCCACTGTTTCAACCGTCTTACCTGTTGCCCTGCTCGATCTTCTTC	600
541	Db	CAGTCTCATCTGGATCCACTGTTTCAACCGTCTTACCTGTTGCCCTGCTCGATCTTCTTC	600
601	Qy	ATCTTGAATCAATCTTTGTGTACAAGCTCAGGAGGAAGACAATTTTGTCTCCGTGGC	660
601	Db	ATCTTGAATCAATCTTTGTGTACAAGCTCAGGAGGAAGACAATTTTGTCTCCGTGGC	660
661	Qy	TACTCCACGGGGAAGACACCGCCATCTTGTGTACCAATTAACCTCCATCTTTGCCACACTT	720
661	Db	TACTCCACGGGGAAGACACCGCCATCTTGTGTACCAATTAACCTCCATCTTTGCCACACTT	720
721	Qy	TGGGCCCCCGCATCATGATTTTTCACACTCTATGAGGGGCGCCATCCAGAACCGC	780
721	Db	TGGGCCCCCGCATCATGATTTTTCACACTCTATGAGGGGCGCCATCCAGAACCGC	780
781	Qy	TGGCTGTGACATCATGTCGGAATTTGCCAAACATGCTAGCCCTTCTGAACACAGCCATC	840
781	Db	TGGCTGTGACATCATGTCGGAATTTGCCAAACATGCTAGCCCTTCTGAACACAGCCATC	840
841	Qy	AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCCACCATGGCAGCGCCACCGCTC	900
841	Db	AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCCACCATGGCAGCGCCACCGCTC	900
901	Qy	AAGCTTTCTTCAAGTGCAGAGCAACCTGTGA CAGTTCTACACCAATCAAACTTTTCC	960
901	Db	AAGCTTTCTTCAAGTGCAGAGCAACCTGTGA CAGTTCTACACCAATCAAACTTTTCC	960
961	Qy	ATAACAGTAGCCCCCTGGATCTCGCCGGCAACCTCACACTGCATCAAGATGCTGGTGTAC	1020
961	Db	ATAACAGTAGCCCCCTGGATCTCGCCGGCAACCTCACACTGCATCAAGATGCTGGTGTAC	1020
1021	Qy	CAGTATGACAAAAATGAAAAACCTATAAAAAGTCGTAAATGACACAGAAAAAGCTCTCTACAG	1080
1021	Db	CAGTATGACAAAAATGAAAAACCTATAAAAAGTCGTAAATGACACAGAAAAAGCTCTCTACAG	1080
1081	Qy	TTTGAAGATGCCATTGGAGCTTGTGTCAATCTCTGTA	1119
1081	Db	TTTGAAGATGCCATTGGAGCTTGTGTCAATCTCTGTA	1119

[illegible]

**REFERENCE**  
**AUTHORS**  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 1.  
 Thornton,M., Patterson,C., Ial,P., Burford,N., Yue,H., Gandhi,A.R.,  
 Elliot,V.S., Ramkumar,J., Baughn,M.R., Kallick,D.A., Walia,N.K.,  
 Hafalia,A.J., Luo,M.G., Lu,Y., Tribouley,C.M., Policky,J.L.,  
 Kearney,L., Graul,R.C., Warren,B.A. and Bing,L.  
**TITLE**  
 G-Protein coupled receptors  
 Patent: WO 0210387-A 25 07-FEB-2002;  
**JOURNAL**  
 Incyte Genomics, Inc. (US)  
**FEATURES**  
 Location/Qualifiers  
 1..1130  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 /note="Incyte ID No: 7474767CB1"

Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

```
QY 781 TGGCTGTGACATCATGTCGCGACATTTGCCAACAATGCTAGCCCTTCTGAAACAGCGCATC 840
Db 792 TGGCTGTGACATCATGTCGCGACATTTGCCAACAATGCTAGCCCTTCTGAAACAGCGCATC 851
QY 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGGACACATGCGCGCGCCAGCGTTC 900
Db 852 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGGACACATGCGCGCGCCAGCGTTC 911
QY 901 AAGGCTTCTTCAAGTGCAGAGCAACCTGTATACAGTTCTTACACCAATCATAACTTTTTC 960
Db 912 AAGGCTTCTTCAAGTGCAGAGCAACCTGTATACAGTTCTTACACCAATCATAACTTTTTC 971
QY 961 ATAAAGTATAGCCCTTGGATCTCGCGGCAAACTCACTCATGATCAAGATGCTGGTGTAC 1020
Db 972 ATAAAGTATAGCCCTTGGATCTCGCGGCAAACTCACTCATGATCAAGATGCTGGTGTAC 1031
QY 1021 CAGTATGACAAAATGGAACCTTATAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 1080
Db 1032 CAGTATGACAAAATGGAACCTTATAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 1091
QY 1081 TTTGAAGATGCCATTTGGAGCTTGTGTCATCATCTCTGTGA 1119
Db 1092 TTTGAAGATGCCATTTGGAGCTTGTGTCATCATCTCTGTGA 1130

RESULT 3
AX463227
LOCUS AX463227 1059 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 1 from Patent WO0248358.
ACCESSION AX463227
VERSION AX463227.1 GI:21886199
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Smolyar,A., Zhu,Z., Encinas,J., Watanabe,S. and Okigami,H.
TITLE Regulation of human chemokine-like receptor
JOURNAL Patent: WO 0248358-A 1 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
Source
1. 1059
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 93.8%; Score 1050; DB 6; Length 1059;
Best Local Similarity 100.0%; Pred. No. 6.5e-252;
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACAGCAGCGCCACCTCGCAGCGCAACAGCTCGCTGCTTGTGGTGTCCCGCGG 60
Db 1 ATGGAGCACAGCAGCGCCACCTCGCAGCGCAACAGCTCGCTGCTTGTGGTGTCCCGCGG 60
QY 61 TCGGCTCGCGCTTGGTGTTCGCGCGGTGCTCTACTACAGCTCTTGTGTCCTCGGT 120
Db 61 TCGGCTCGCGCTTGGTGTTCGCGCGGTGCTCTACTACAGCTCTTGTGTCCTCGGT 120
QY 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 180
Db 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 180
QY 181 TCCTCTCAAACTATCTTTGGCACTGCTGCTGCGGCACTATGCTGCTCTTTTTCATA 240
Db 181 TCCTCTCAAACTATCTTTGGCACTGCTGCTGCGGCACTATGCTGCTCTTTTTCATA 240
QY 241 GTCTTTGTGACTTCTGTTGGAAGTTTCATCTTGAACATGAGATGCTCAGGTCCCC 300
Db 241 GTCTTTGTGACTTCTGTTGGAAGTTTCATCTTGAACATGAGATGCTCAGGTCCCC 300
QY 301 GACAAGATCATAGAAGTGTGGAATTCTCATCCATCCACACCTCCATATGATTACTGTA 360
```

```
Db 301 GACAAGATCATAGAAGTGTGGAATTCTCATCCATCCACACCTCCATATGATTACTGTA 360
QY 361 CGGTTAAACCATTTGACAGGTATATCGCTGCTCGCCACCGCTCAAGTACACACGCTCTCA 420
Db 361 CGGTTAAACCATTTGACAGGTATATCGCTGCTCGCCACCGCTCAAGTACACACGCTCTCA 420
QY 421 TACCAGCCCGCACCGGAAAGTCAATTGTAAGTGTTTATACATCACCTGCTTCTTGACACAG 480
Db 421 TACCAGCCCGCACCGGAAAGTCAATTGTAAGTGTTTATACATCACCTGCTTCTTGACACAG 480
QY 481 ATCCCTTATTTACTGTTGGCCCAACATCTGGACTGAAGACTATACAGCACTCTGTGTCAT 540
Db 481 ATCCCTTATTTACTGTTGGCCCAACATCTGGACTGAAGACTATACAGCACTCTGTGTCAT 540
QY 541 CAGGCTCTCATCTGGATCCACTGCTTACCGCTTACCTGCTGCTGCTGCTTCTTCTTCTTC 600
Db 541 CAGGCTCTCATCTGGATCCACTGCTTACCGCTTACCTGCTGCTGCTGCTTCTTCTTCTTC 600
QY 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTCTCTCCGTGGC 660
Db 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTCTCTCCGTGGC 660
QY 661 TACTCCACGGGGAAGACACCGCCCATCTTGTTCACCATTTACCTTCTTGTCCACACTT 720
Db 661 TACTCCACGGGGAAGACACCGCCCATCTTGTTCACCATTTACCTTCTTGTCCACACTT 720
QY 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGGGGCGCCCATCCAGAACCCG 780
Db 721 TGGGCCCCCGCATCATCATGATTTCTTACCACTCTATGGGGCGCCCATCCAGAACCCG 780
QY 781 TGGCTGGTACACATCATGTCCGACATTTGCCAATGCTAGCCCTTCTGAAACAGCGCATC 840
Db 781 TGGCTGGTACACATCATGTCCGACATTTGCCAATGCTAGCCCTTCTGAAACAGCGCATC 840
QY 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGGACACATGCGCAGCGCCAGCTC 900
Db 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGGACACATGCGCAGCGCCAGCTC 900
QY 901 AAGGCTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCATAACTTTTTC 960
Db 901 AAGGCTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTTACCACTCATAACTTTTTC 960
QY 961 ATAAAGTATAGCCCTTGGATCTCGCGGCAAACTCACTGATCAAGATGCTGGTGTAC 1020
Db 961 ATAAAGTATAGCCCTTGGATCTCGCGGCAAACTCACTGATCAAGATGCTGGTGTAC 1020
QY 1021 CAGTATGACAAAATGGAACCTTATAA 1050
Db 1021 CAGTATGACAAAATGGAACCTTATAA 1050

RESULT 4
BD183203
LOCUS BD183203
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD183203
VERSION BD183203.1 GI:31875403
KEYWORDS JP 2002345481-A/3.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 1062)
AUTHORS Terao,Y., Matsui,H. and Shintani,Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: JP 2002345481-A 3 03-DEC-2002;
TAKEDA CHEMICAL INDUSTRIES LTD
COMMENT
OS Homo sapiens (human)
PN JP 2002345481-A/3
PD 03-DEC-2002
PF 01-JUN-2001 JP 2001166688
PI YASUKO TERA0,HIDEKI MATSUI,YASUSHI SHINTANI
```

PC C12N15/09,A61K45/00,A61P1/00,A61P3/00,A61P9/00,A61P25/00, PC  
A61P29/00,  
PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC  
G01N33/53,  
PC G01N33/566,C12N15/00,C12N5/00  
CC Novel G protein-coupled receptor and its DNA  
FH Key Location/Qualifiers  
FT source 1..1062  
FT /organism='Homo sapiens (human)'.  
FEATURES  
source  
Location/Qualifiers  
1..1062  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'  
ORIGIN  
Query Match 93.8%; Score 1050; DB 6; Length 1062;  
Best Local Similarity 100.0%; Pred. No. 6.5e-252;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACACGACGCCACCTCGCAGCCAAAGAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
Db 1 ATGGAGCACACGACGCCACCTCGCAGCCAAAGAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
QY 61 TCGGCTCGCGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTGCTCGGT 120  
Db 61 TCGGCTCGCGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTGCTCGGT 120  
QY 121 TTACGAGCAATATCTTGACAGTGATCATCCTCTCCAGCTGGTGGCAAGACAGAG 180  
Db 121 TTACGAGCAATATCTTGACAGTGATCATCCTCTCCAGCTGGTGGCAAGACAGAG 180  
QY 181 TCTCCTACAATATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA 240  
Db 181 TCTCCTACAATATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA 240  
QY 241 GTGTTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCTCAGTCCGC 300  
Db 241 GTGTTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCTCAGTCCGC 300  
QY 301 GACAGATCATAGAAGTGTGGAAATCTCATCATCCATCCACCTCCATATGGAATTAAGTGA 360  
Db 301 GACAGATCATAGAAGTGTGGAAATCTCATCATCCATCCACCTCCATATGGAATTAAGTGA 360  
QY 361 CGGTTAACCATTTGACAGGTTATATCGCTGTGCGCACCGCTCAAGTACCAACAGGTTCA 420  
Db 361 CGGTTAACCATTTGACAGGTTATATCGCTGTGCGCACCGCTCAAGTACCAACAGGTTCA 420  
QY 421 TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTTCATCATCCTGCTTCTTGACCAAGC 480  
Db 421 TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTTCATCATCCTGCTTCTTGACCAAGC 480  
QY 481 ATCCCTCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGAT 540  
Db 481 ATCCCTCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGAT 540  
QY 541 CAGCTCCTCATCTGGATCCATCGCTTTCACCGTCTACCTGTGGTCCCTGTCTCACTTTCTTC 600  
Db 541 CAGCTCCTCATCTGGATCCATCGCTTTCACCGTCTACCTGTGGTCCCTGTCTCACTTTCTTC 600  
QY 601 ATCTTGAATCTCAATCTGTGTGACAGCTCAGAGGAGAGCAATTTTGGTCTCGTGGC 660  
Db 601 ATCTTGAATCTCAATCTGTGTGACAGCTCAGAGGAGAGCAATTTTGGTCTCGTGGC 660  
QY 661 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACATTTACCTCTTTTGGCCACATTT 720  
Db 661 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACATTTACCTCTTTTGGCCACATTT 720  
QY 721 TGGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 721 TGGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780

QY 781 TGGCTGGTACACATCATCTCGACATTCGCCAACATGTCAGCCCTTCTTGAAACACAGCCATC 840  
Db 781 TGGCTGGTACACATCATCTCGACATTCGCCAACATGTCAGCCCTTCTTGAAACACAGCCATC 840  
QY 841 AACTTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCACCATGCGCAGCGCCACGCTC 900  
Db 841 AACTTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCACCATGCGCAGCGCCACGCTC 900  
QY 901 AAGGCTTTCTTCAAGTGCCAGAACCTGTACAGTTCTACAGTTCTACCAATCATAACTTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGCCAGAACCTGTACAGTTCTACCAATCATAACTTTTCC 960  
QY 961 ATAAACAAGTAGCCCTGGATCTCGCGGCAAACTCACAATCAAGATCTGTGTGTAC 1020  
Db 961 ATAAACAAGTAGCCCTGGATCTCGCGGCAAACTCACAATCAAGATCTGTGTGTAC 1020  
QY 1021 CAGTATGACAAAAATGAAAACTATATAA 1050  
Db 1021 CAGTATGACAAAAATGAAAACTATATAA 1050  
RESULT 5  
BD103851  
LOCUS Novel G protein-coupled receptor and its DNA. linear PAT 27-AUG-2002  
DEFINITION Novel G protein-coupled receptor and its DNA.  
ACCESSION BD103851  
VERSION BD103851.1 GI:22649425  
KEYWORDS WO 0194582-A/3.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1062)  
AUTHORS Terao,Y., Matsui,H. and Shintani,Y.  
TITLE Novel G protein-coupled receptor and its DNA  
JOURNAL Patent: WO 0194582-A 3 13-DEC-2001;  
TAKEDA CHEMICAL INDUSTRIES LTD,YASUKO TERAO,HIDEKI MATSUI,YASUSHI  
SHINTANI  
COMMENT OS Homo sapiens (human)  
PN WO 0194582-A/3  
PD 13-DEC-2001  
PF 01-JUN-2001 WO 2001JP004643  
PR 02-JUN-2000 JP 00P 170446, 23-JUN-2000 JP 00P 194926 PI  
YASUKO TERAO,HIDEKI MATSUI,YASUSHI SHINTANI  
PC C12N15/12,C12P21/02,C07K14/705,C07K16/28,A61K45/00,A61P25/00,  
PC A61P29/00,  
PC A61P9/00,A61P35/00,A61P37/00,A61P1/00,A61P3/50,  
G01N33/15  
PC G01N33/15  
CC Novel G protein-coupled receptor and its DNA  
FH Key Location/Qualifiers  
FT source 1..1062  
FT /organism='Homo sapiens (human)'.  
FEATURES  
source  
Location/Qualifiers  
1..1062  
/organism='Homo sapiens'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:9606'  
ORIGIN  
Query Match 93.8%; Score 1050; DB 6; Length 1062;  
Best Local Similarity 100.0%; Pred. No. 6.5e-252;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGAGACACAGCAGCCACCTCGCAGCCAAAGAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
Db 1 ATGAGACACAGCAGCCACCTCGCAGCCAAAGAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
QY 61 TCGGCTCGCGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTGCTCGGT 120  
Db 61 TCGGCTCGCGCTTGGGTTTCGTGCGCGGTCTACTACAGCCTCTTGTGTGCTCGGT 120  
QY 121 TTACGAGCAATATCTTGACAGTGATCATCCTCTCCAGCTGGTGGCAAGACAGAG 180  
Db 121 TTACGAGCAATATCTTGACAGTGATCATCCTCTCCAGCTGGTGGCAAGACAGAG 180  
QY 181 TCTCCTACAATATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA 240  
Db 181 TCTCCTACAATATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA 240  
QY 241 GTGTTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCTCAGTCCGC 300  
Db 241 GTGTTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCTCAGTCCGC 300  
QY 301 GACAGATCATAGAAGTGTGGAAATCTCATCATCCATCCACCTCCATATGGAATTAAGTGA 360  
Db 301 GACAGATCATAGAAGTGTGGAAATCTCATCATCCATCCACCTCCATATGGAATTAAGTGA 360  
QY 361 CGGTTAACCATTTGACAGGTTATATCGCTGTGCGCACCGCTCAAGTACCAACAGGTTCA 420  
Db 361 CGGTTAACCATTTGACAGGTTATATCGCTGTGCGCACCGCTCAAGTACCAACAGGTTCA 420  
QY 421 TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTTCATCATCCTGCTTCTTGACCAAGC 480  
Db 421 TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTTCATCATCCTGCTTCTTGACCAAGC 480  
QY 481 ATCCCTCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGAT 540  
Db 481 ATCCCTCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGAT 540  
QY 541 CAGCTCCTCATCTGGATCCATCGCTTTCACCGTCTACCTGTGGTCCCTGTCTCACTTTCTTC 600  
Db 541 CAGCTCCTCATCTGGATCCATCGCTTTCACCGTCTACCTGTGGTCCCTGTCTCACTTTCTTC 600  
QY 601 ATCTTGAATCTCAATCTGTGTGACAGCTCAGAGGAGAGCAATTTTGGTCTCGTGGC 660  
Db 601 ATCTTGAATCTCAATCTGTGTGACAGCTCAGAGGAGAGCAATTTTGGTCTCGTGGC 660  
QY 661 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACATTTACCTCTTTTGGCCACATTT 720  
Db 661 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACATTTACCTCTTTTGGCCACATTT 720  
QY 721 TGGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 721 TGGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780

```
QY 121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGCTGGCAAGACAGAG 180
Db 121 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGCTGGCAAGACAGAG 180
QY 181 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGGCACTCTTGGCTCTTTTTCATA 240
Db 181 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGGCACTCTTGGCTCTTTTTCATA 240
QY 241 GTGTTTGTGCACTCTCTTGGCAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCC 300
Db 241 GTGTTTGTGCACTCTCTTGGCAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCC 300
QY 301 GACAAGATCATAGAGTCTGGAAATTTCTCATCTCCACCTCCATATGATTTACTGTA 360
Db 301 GACAAGATCATAGAGTCTGGAAATTTCTCATCTCCACCTCCATATGATTTACTGTA 360
QY 361 CCGTTAAACCATGACAGTATATCGCTGCTGCGACCGCTCAAGTACACACAGGTCTCA 420
Db 361 CCGTTAAACCATGACAGTATATCGCTGCTGCGACCGCTCAAGTACACACAGGTCTCA 420
QY 421 TACCCAGCCGCAACCCGGAAGTCAATGTAAGTGTTCATCATCTGCTTCTTGACCAGC 480
Db 421 TACCCAGCCGCAACCCGGAAGTCAATGTAAGTGTTCATCATCTGCTTCTTGACCAGC 480
QY 481 ATCCCTTATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
Db 481 ATCCCTTATTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540
QY 541 CAGTCTCTCATCTGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
Db 541 CAGTCTCTCATCTGGAATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 601 ATCTTGAACATCAATCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 660
Db 601 ATCTTGAACATCAATCTGTAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 660
QY 661 TACTCCAGGGGGAAGACACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
Db 661 TACTCCAGGGGGAAGACACCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
QY 721 TGGGCCCCCGCATCATGATTTTACCACTCTATGGGGGCGCCATCCAGAACCGC 780
Db 721 TGGGCCCCCGCATCATGATTTTACCACTCTATGGGGGCGCCATCCAGAACCGC 780
QY 781 TGGTGGTACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGACATGTC 840
Db 781 TGGTGGTACATCATGTCGACATGTCGACATGTCGACATGTCGACATGTCGACATGTC 840
QY 841 AACTTCTCTCTACTGCTTATCAGAGCGGTTGCGACACATGAGCGGCGCCAGCTC 900
Db 841 AACTTCTCTCTACTGCTTATCAGAGCGGTTGCGACACATGAGCGGCGCCAGCTC 900
QY 901 AAGGCTTCTTCAAGTGCAGAGCACTGTCAGTCTTACACCAATCATATCTTTTCC 960
Db 901 AAGGCTTCTTCAAGTGCAGAGCACTGTCAGTCTTACACCAATCATATCTTTTCC 960
QY 961 ATAAAGTAGCCCTGGAATCTGCGCGGAACTCAGCTGCACTGCACTGCTGCTGCTGCT 1020
Db 961 ATAAAGTAGCCCTGGAATCTGCGCGGAACTCAGCTGCACTGCACTGCTGCTGCTGCT 1020
QY 1021 CAGTATGACAAAATGGAACCTATAAAA 1050
Db 1021 CAGTATGACAAAATGGAACCTATAAAA 1050
```

```
RESULT 6
AY635179 1167 bp mRNA linear PRI 04-JUL-2004
LOCUS Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete
DEFINITION cds.
ACCESSION AY635179
VERSION AY635179.1 GI:49413551
KEYWORDS
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 1167)
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Complete coding sequence of GPR139
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1167)
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06,
MSC4094, Bethesda, MD 20892-4094, USA
FEATURES
source
1..1167
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosomes="16"
/map="16p12.3"
/tissue_type="brain"
1..1167
/gene="GPR139"
92..1153
/gene="GPR139"
/codon_start=1
/product="G protein-coupled receptor 139"
/protein_id="AAT65818.1"
/db_xref="GI:49413552"
/translation="MEHTAHLAANSSLSWSPGSCAGLGFVVPVYVYLLCGLPAN
IUTVILSQLVARRKSSYNLLAADAALVFFVDFVDFLEDFILNQMFQVDPK
IIEVLEFSIHSTIWITVPLTDITDITVAVCHPLKHTVSYPARTKRVISYVITCFPTS
IPYWPNIWEDYISTSVHVLIMHCTVTVLPCSIFFILNSIIVYKLRKSNFRL
RGVSTKTAIFTTSTISFATLWAPRIIMLVHLYGAPIONRWLHIMSDIANMLALL
NTAINFLYCFISKBFRTMAATLKAFFKQKQVQPYTNHFSITSSPWISPAHSCH
IKMLVYQDKNKPKIVSP"
primer_bind complement(1142..1167)
Query Match 93.8%; Score 1049.6; DB 9; Length 1167;
Best Local Similarity 99.9%; Pred. No. 8.3e-252;
Matches 1049; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGGAGCACACAGCAGCCACCTCGCAGCAACAGCTCGCTGTGTGGTGGTCCCCGGC 60
Db 92 ATGGAGCACACAGCAGCCACCTCGCAGCAACAGCTCGCTGTGTGGTGGTCCCCGGC 151
QY 61 TGGGCTTGGGCTTGGGTTTCTGCGCGTGGTCTACTACAGCTCTTGTGTGCTCGGT 120
Db 152 TGGGCTTGGGCTTGGGTTTCTGCGCGTGGTCTACTACAGCTCTTGTGTGCTCGGT 211
QY 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGCTGGCAAGACAGAG 180
Db 212 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGCTGGCAAGACAGAG 271
QY 181 TCCTCTCAAACTATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 240
Db 272 TCCTCTCAAACTATCTCTTGGCACTCGCTGCTGCGGACATCTTGGTCTCTTTTTCATA 331
QY 241 GTGTTTGTGCACTCTCTTGGCAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCC 300
Db 332 GTGTTTGTGCACTCTCTTGGCAAGATTTTCATCTTGAACATGAGATGCTCAGGTCCCC 391
QY 301 GACAAGATCATAGAGTCTGGAAATTTCTCATCTCCACCTCCATATGATTTACTGTA 360
Db 392 GACAAGATCATAGAGTCTGGAAATTTCTCATCTCCACCTCCATATGATTTACTGTA 451
QY 361 CGGTTAAACCATGACAGGTATATCGCTGCTGCGACCGCTCAAGTACACACAGGTCTCA 420
Db 452 CGGTTAAACCATGACAGGTATATCGCTGCTGCGACCGCTCAAGTACACACAGGTCTCA 511
QY 421 TACCCAGCCGCAACCCGGAAGTCAATGTAAGTGTTCATCATCTGCTTCTTGACCAGC 480
```

Db 512 TACCCAGCCGACCGGAAAGTCATTGTGAAGTGTTCATCACCTGCTCTCTGACCAGC 571  
QY 481 ATCCCTCTATTACTGTGGGCCCAACATCTGAGCTGAAGACTATACAGCACTCTGTGCAT 540  
Db 572 ATCCCTCTATTACTGTGGGCCCAACATCTGAGCTGAAGACTATACAGCACTCTGTGCAT 631  
QY 541 CAGCTCCTCATCTGGATCCACTGCTTACCGGTCTACCTGGTGCCTGCTCCATCTTCTTC 600  
Db 632 CAGCTCCTCATCTGGATCCACTGCTTACCGGTCTACCTGGTGCCTGCTCCATCTTCTTC 691  
QY 601 ATCTTGAATCAATATGTTGATACAGCTCAGAGGAAGACAAATTTGCTCCGTGCG 660  
Db 692 ATCTTGAATCAATATGTTGATACAGCTCAGAGGAAGACAAATTTGCTCCGTGCG 751  
QY 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACCATTTACCTCTTTCGCACTT 720  
Db 752 TACTCCAGGGGAGACACCGCCATCTTGTTCACCATTTACCTCTTTCGCACTT 811  
QY 721 TGGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 812 TGGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 871  
QY 781 TGGCTGGTACACATCATGTCGCAATTCGCAATGTCATGTCAGCCCTTCTGAACACAGCCATC 840  
Db 872 TGGCTGGTACACATCATGTCGCAATTCGCAATGTCATGTCAGCCCTTCTGAACACAGCCATC 931  
QY 841 NACTTCTTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGGCAGCGCCAGGCTC 900  
Db 932 NACTTCTTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGGCAGCGCCAGGCTC 991  
QY 901 AAGGCTTTCTTCAAGTGCCAGAACCACTGTACAGTTCTTACCACTATATACTTTTCC 960  
Db 992 AAGGCTTTCTTCAAGTGCCAGAACCACTGTACAGTTCTTACCACTATATACTTTTCC 1051  
QY 961 ATAAAGTAGCCCTTGGATCTGCGCGCAAACTCACACTGATCAAGATGCTGCTGTAC 1020  
Db 1052 ATAAAGTAGCCCTTGGATCTGCGCGCAAACTCACACTGATCAAGATGCTGCTGTAC 1111  
QY 1021 CAGTATGACAAAATGGAACCTATAAA 1050  
Db 1112 CAGTATGACAAAATGGAACCTATAAA 1141

RESULT 7  
BD183204  
LOCUS BD183204  
DEFINITION Novel G protein-coupled receptor and its DNA.  
ACCESSION BD183204  
VERSION BD183204.1 GI:31875404  
KEYWORDS JP 2002345481-A/4.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
Terao,Y., Matsui,H. and Shintani,Y.  
1 (bases 1 to 1062)  
Novel G protein-coupled receptor and its DNA  
Patent: JP 2002345481-A 4 03-DEC-2002;  
TAKEDA CHEMICAL INDUSTRIES LTD  
OS Homo sapiens (human)  
PN JP 2002345481-A/4  
PD 03-DEC-2002  
PF 01-JUN-2001 JP 2001166688  
PI YASUKO TERAO,HIDEKI MATSUI,YASUSHI SHINTANI  
PC C12N15/09,A61K45/00,A61P1/00,A61P3/00,A61P9/00,A61P25/00, PC  
A61P29/00,  
PC A61P35/00,A61P37/00,C07K14/705,C07K16/28,C12N1/15,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC  
G01N33/53,  
PC G01N33/566,C12N15/00,C12N5/00  
CC Novel G protein-coupled receptor and its DNA  
FH Key Location/Qualifiers

FT source 1..1062  
FT Location/Qualifiers  
1..1062  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 93.7%; Score 1048.4; DB 6; Length 1062;  
Best Local Similarity 99.9%; Pred. No. 1.6e-251;  
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGGAGCACGCGCCACCTCGCAGCCACAGCTCGCTGCTGTGGTGGTCCCGCGC 60  
Db 1 ATGGAGCACGCGCCACCTCGCAGCCACAGCTCGCTGCTGTGGTGGTCCCGCGC 60  
QY 61 TCGGCTCGGCTTGGGTTTCGTCGCCGTGGTCTACTACAGCCTCTGTGTCCTCGT 120  
Db 61 TCGGCTCGGCTTGGGTTTCGTCGCCGTGGTCTACTACAGCCTCTGTGTCCTCGT 120  
QY 121 TTACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
Db 121 TTACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
QY 181 TCCTCTACAATATCTTGGCACTCGCTGTCGCCGACATCTTGGTCTCTTTTTCATA 240  
Db 181 TCCTCTACAATATCTTGGCACTCGCTGTCGCCGACATCTTGGTCTCTTTTTCATA 240  
QY 241 GTGTTGTGGACTTCTCTTGGAAATTTTCATCTTGAACATGCGCTCAGTCCCGC 300  
Db 241 GTGTTGTGGACTTCTCTTGGAAATTTTCATCTTGAACATGCGCTCAGTCCCGC 300  
QY 301 GACAAGATCATAGAAGTCTGGAATTTCTATCCATCCACACCTCATATGATTAATGTA 360  
Db 301 GACAAGATCATAGAAGTCTGGAATTTCTATCCATCCACACCTCATATGATTAATGTA 360  
QY 361 CGGTTAACCATTTGACAGGTATATCGCTGTGCGCACCGCTCAAGTACACACGGTCTCA 420  
Db 361 CGGTTAACCATTTGACAGGTATATCGCTGTGCGCACCGCTCAAGTACACACGGTCTCA 420  
QY 421 TACCCAGCCGACCCCGGAAAGTCAATTGTAAGTGTTCATCACCTGCTTCTGACCAGC 480  
Db 421 TACCCAGCCGACCCCGGAAAGTCAATTGTAAGTGTTCATCACCTGCTTCTGACCAGC 480  
QY 481 ATCCCTCTATTACTGTGGGCCCAACATCTGAGCTGAAGACTATACAGCACTCTGTGCAT 540  
Db 481 ATCCCTCTATTACTGTGGGCCCAACATCTGAGCTGAAGACTATACAGCACTCTGTGCAT 540  
QY 541 CAGCTCCTCATCTGGATCCACTGCTTACCGCTTACCTGCTGCTGCTGCTCCTCTTTC 600  
Db 541 CAGCTCCTCATCTGGATCCACTGCTTACCGCTTACCTGCTGCTGCTCCTCTTTC 600  
QY 601 ATCTTGAATCAATATGTTGATACAGCTCAGAGGAAGACAAATTTGCTCCGTGCG 660  
Db 601 ATCTTGAATCAATATGTTGATACAGCTCAGAGGAAGACAAATTTGCTCCGTGCG 660  
QY 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACCATTTACCTCATCTTTCGCACTT 720  
Db 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACCATTTACCTCATCTTTCGCACTT 720  
QY 721 TGGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 721 TGGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
QY 781 TGGCTGGTACACATCATGTCGCAATTCGCAATGTCATGTCAGCCCTTCTGAACACAGCCATC 840  
Db 781 TGGCTGGTACACATCATGTCGCAATTCGCAATGTCATGTCAGCCCTTCTGAACACAGCCATC 840  
QY 841 NACTTCTTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGGCAGCGCCAGGCTC 900  
Db 841 NACTTCTTCTCTACTGCTTCATCAGCAGCGGTTCCGACCATGGCAGCGCCAGGCTC 900



```
QY 901 AAGCTTTCTTCAAGTCCAGAGACCACTGTACAGTTCTACACCAATCATACTTTTC 960
Db 901 AAGCTTTCTTCAAGTCCAGAGACCACTGTACAGTTCTACACCAATCATACTTTTC 960
QY 961 ATAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACATGCAATCAAGATGCTGGGTGAC 1020
Db 961 ATAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACATGCAATCAAGATGCTGGGTGAC 1020
QY 1021 CAGTATGACAAAATGGAACCTTATAAA 1050
Db 1021 CAGTATGACAAAATGGAACCTTATAAA 1050

RESULT 8
AX498192
LOCUS AX498192 1062 bp DNA linear PAT 26-SBP-2002
DEFINITION Sequence 15 from Patent WO242461.
ACCESSION AX498192
VERSION AX498192.1 GI:23343117
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1
AUTHORS Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
TITLE Endogenous and non-endogenous versions of human g protein-coupled
JOURNAL receptors
Patent: WO 0242461-A 15 30-MAY-2002;
Arena Pharmaceuticals, Inc. (US)
FEATURES
source Location/Qualifiers
1..1062
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="Novel Sequence"
ORIGIN
Query Match 93.7%; Score 1048.4; DB 6; Length 1062;
Best Local Similarity 99.9%; Pred. No. 1.6e-251;
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGGAGCAGCAGCAGCCACCTCGCAGCAACAGCTCGCTGCTTGGTGGTCCCGGC 60
Db 1 ATGGAGCAGCAGCAGCCACCTCGCAGCAACAGCTCGCTGCTTGGTGGTCCCGGC 60
QY 61 TCGGCTCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
Db 61 TCGGCTCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 120
QY 121 TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180
Db 121 TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180
QY 181 TCCTCTACAACTATCTTGGCAGCTCGCTGCTGCGGACATCTTGGTCTCTTTTCATA 240
Db 181 TCCTCTACAACTATCTTGGCAGCTCGCTGCTGCGGACATCTTGGTCTCTTTTCATA 240
QY 241 GTGTTGTGAGCTTCTGTTGGAGATTTTCATCTTGAAATGCAATGCAATGCTCAGGTCCCC 300
Db 241 GTGTTGTGAGCTTCTGTTGGAGATTTTCATCTTGAAATGCAATGCAATGCTCAGGTCCCC 300
QY 301 GACAAGATCATAGAAGTGTGGAATTTCTATCATCCACACCTCCATATGATGATTACTGTA 360
Db 301 GACAAGATCATAGAAGTGTGGAATTTCTATCATCCACACCTCCATATGATGATTACTGTA 360
QY 361 CGGTTAACCAATTGACAGGTATATCGCTGTCTGCGACCCCGCTCAAGTACCAACAGGTCTCA 420
Db 361 CGGTTAACCAATTGACAGGTATATCGCTGTCTGCGACCCCGCTCAAGTACCAACAGGTCTCA 420
QY 421 TACCCAGCCGCGACCCGGAAGTCATTGTAAGTGTATACATCACCTGCTTCTGACGACG 480
Db 421 TACCCAGCCGCGACCCGGAAGTCATTGTAAGTGTATACATCACCTGCTTCTGACGACG 480
```

```
QY 481 ATCCCCCTATTACTGGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGGCAT 540
Db 481 ATCCCCCTATTACTGGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGGCAT 540
QY 541 CACGTCCTCATCTGGATCCACTGCTTACCGTCTACCTGCTGCGCTGCTGCTTCTTCTTC 600
Db 541 CACGTCCTCATCTGGATCCACTGCTTACCGTCTACCTGCTGCGCTGCTGCTTCTTCTTC 600
QY 601 ATCTTGAACCTCAATCTGTGTACAGCTCAGAGGGAAGCAATTTTCTCTCCGTGGC 660
Db 601 ATCTTGAACCTCAATCTGTGTACAGCTCAGAGGGAAGCAATTTTCTCTCCGTGGC 660
QY 661 TACTCCACGGGGAAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTGGCCACAT 720
Db 661 TACTCCACGGGGAAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTGGCCACAT 720
QY 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCOC 780
Db 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCOC 780
QY 781 TGGCTGGTACACATCATGTCCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 840
Db 781 TGGCTGGTGGTACATCATGTCCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 840
QY 841 AACTTTCTCTCTACTGCTTCTATCAGCAAGCGGTTCCGACCAATGCGCAGCGCCAGCTC 900
Db 841 AACTTTCTCTCTACTGCTTCTATCAGCAAGCGGTTCCGACCAATGCGCAGCGCCAGCTC 900
QY 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTCTTACACCAATCATACTTTTCC 960
Db 901 AAGGCTTTCTTCAAGTGGCCAGAGCAACCTGTACAGTCTTACACCAATCATACTTTTCC 960
QY 961 ATAAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGCAATCAAGATGCTGGTGTAC 1020
Db 961 ATAAACAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGCAATCAAGATGCTGGTGTAC 1020
QY 1021 CAGTATGACAAAATGGAACCTTATAAA 1050
Db 1021 CAGTATGACAAAATGGAACCTTATAAA 1050

RESULT 9
BD103852
LOCUS BD103852 1062 bp DNA linear PAT 27-AUG-2002
DEFINITION Novel G protein-coupled receptor and its DNA.
ACCESSION BD103852
VERSION BD103852.1 GI:22649426
KEYWORDS WO 0194582-A/4.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1062)
Terao, Y., Matsui, H. and Shintani, Y.
Novel G protein-coupled receptor and its DNA
Patent: WO 0194582-A 4 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI
SHINTANI
OS Homo sapiens (human)
PN WO 0194582-A/4
PD 13-DEC-2001
PR 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP 00P 170446.23-JUN-2000 JP 00P 194926 PI
YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI
PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00,
PC A61P29/00,
PC A61P09/00, A61P35/00, A61P37/00, A61P1/00, G01N33/566, PC
G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
/organism='Homo sapiens (human)'.
FT
```



Db	482	ATCCCCCTATTACTGGTGGCCCAACATCTGGAAGAGACTATACGACACTCTGTGCAT	541
Qy	541	CAGTCTCTCATCTGGATFCACTGTTTCCACCGTCTACCTGGTGGCCCTGCTCCATCTTCTTC	600
Db	542	CAGTCTCTCATCTGGATFCACTGTTTCCACCGTCTACCTGGTGGCCCTGCTCCATCTTCTTC	601
Qy	601	ATCTTGAAGTCAATCATCTGTTGATCAAGCTCAGGAGGAGAGCAATTTGCTCTCCGTGGC	660
Db	602	ATCTTGAAGTCAATCATCTGTTGATCAAGCTCAGGAGGAGAGCAATTTGCTCTCCGTGGC	661
Qy	661	TACTTCCACGGGAGACACCGCATCTTGTTCACCATTAACCTCCATCTTTTGCCACACTT	720
Db	662	TACTTCCACGGGAGACACCGCATCTTGTTCACCATTAACCTCCATCTTTTGCCACACTT	721
Qy	721	TGGGCCCCCGCATCATCATGATTTTACCACCTCTATGGGGGCGCCATCCAGAACCCGC	780
Db	722	TGGGCCCCCGCATCATCATGATTTTACCACCTCTATGGGGGCGCCATCCAGAACCCGC	781
Qy	781	TGCTGTGTACACATCATCTGCGACATTTGCCAATGCTAGCCCTTCTGAAACAGCCATC	840
Db	782	TGCTGTGTGTACACATCATCTGCGACATTTGCCAATGCTAGCCCTTCTGAAACAGCCATC	841
Qy	841	AACCTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGAGCGCCGACGCTC	900
Db	842	AACCTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGAGCGCCGACGCTC	901
Qy	901	AAGGCTTTCTTCAAGTGCAGAGAACCTGTGATCTTACACCAATCATATACTTTTTC	960
Db	902	AAGGCTTTCTTCAAGTGCAGAGAACCTGTGATCTTACACCAATCATATACTTTTTC	961
Qy	961	ATAACAGTATAGCCCTGGATCTCGCGGCAAACTCAACATGATCAAGATGCTGGTGTAC	1020
Db	962	ATAACAGTATAGCCCTGGATCTCGCGGCAAACTCAACATGATCAAGATGCTGGTGTAC	1021
Qy	1021	CAGTATGACAAAATGGAACCTTATAAA	1050
Db	1022	CAGTATGACAAAATGGAACCTTATAAA	1051
RESULT 11			
LOCUS	CQ867782 2117 bp DNA linear PAT 13-SEP-2004		
DEFINITION	Sequence 1 from Patent WO2004074841.		
ACCESSION	CQ867782		
VERSION	CQ867782.1 GI:51997901		
KEYWORDS	Homo sapiens (human)		
SOURCE	Homo sapiens		
ORGANISM	Homo sapiens		
REFERENCE	Murphy, A.J. and Croll-Kalish, S.		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
JOURNAL	Murphy, A.J. and Croll-Kalish, S.		
FEATURES	Kor3l-like proteins and methods of modulating kor3l-mediated activity		
source	Patent: WO 2004074841-A 1 02-SEP-2004; REGENERON PHARMACEUTICALS, INC. (US); Murphy, Andrew J. (US); Croll-Kalish, Susan (US) Location/Qualifiers 1. .2117 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606"		
ORIGIN			
Query Match 93.7%; Score 1048.4; DB 6; Length 2117;			
Best Local Similarity 99.9%; Pred. No. 1.8e-251;			
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;			
Qy	1	ATGAGCAGACAGCCGACCTCGAGCAACAGCTCGCTGCTTGGTGGTCCCGCGC	60
Db	280	ATGAGCAGACAGCCGACCTCGAGCAACAGCTCGCTGCTTGGTGGTCCCGCGC	339
Qy	61	TCGGCCTGCGGCTTGGTTCGTGGCGGCTGCTACTACAGCCTCTTGTGCTCGCTCGGT	120

RESULT 12  
AX664703  
LOCUS  
DEFINITION  
Sequence 12 from Patent WO02074960.  
ACCESSION  
AX664703

AX664703  
Sequence 12 from Patent WO02074960.  
linear  
PAT 22-MAR-2003

Db	340	TCGGCCTGCGGCTTGGGTTTCGGTCCCGTGGTCTACTACAGCCTCTTGTGTGCTCGGT	399
Qy	121	TTACCAAGCAAAATATCTTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG	180
Db	400	TTACCAAGCAAAATATCTTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG	459
Qy	181	TGCTCTCAACAATATCTTTGGCACTCGCTGTGTCGCGACATCTTTGGTCTCTTTTTCATA	240
Db	460	TGCTCTCAACAATATCTTTGGCACTCGCTGTGTCGCGACATCTTTGGTCTCTTTTTCATA	519
Qy	241	GTGTTTGTGGACTTCTGTTTGGAGATTTTCACTTTGAAATGCAATGCAAGTGGCTCAGGTCCCC	300
Db	520	GTGTTTGTGGACTTCTGTTTGGAGATTTTCACTTTGAAATGCAATGCAAGTGGCTCAGGTCCCC	579
Qy	301	GACCAAGATCATAGAAGTGTGGAATTTCTATCCATCCACCTCCATATATGGATTTACTGTA	360
Db	580	GACCAAGATCATAGAAGTGTGGAATTTCTATCCATCCACCTCCATATATGGATTTACTGTA	639
Qy	361	CGGTTAACCAATTGACAGGTATATCGCTGTCTGCGACCCGCTCAAGTACCAACAGGCTCTCA	420
Db	640	CGGTTAACCAATTGACAGGTATATCGCTGTCTGCGACCCGCTCAAGTACCAACAGGCTCTCA	699
Qy	421	TACCCAGCCGCGACCCGGAAGTCAATGTAAGTGTGTTTACATCACTGCTCTTCTGACGAGC	480
Db	700	TACCCAGCCGCGACCCGGAAGTCAATGTAAGTGTGTTTACATCACTGCTCTTCTGACGAGC	759
Qy	481	ATCCCTCTATTTACTTGGTGGCCCAACATCTGAGCTTGAAGTGTGTTTACATCACTGCTCTT	540
Db	760	ATCCCTCTATTTACTTGGTGGCCCAACATCTGAGCTTGAAGTGTGTTTACATCACTGCTCTT	819
Qy	541	CAGCTCCTCATCTGGATCCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	600
Db	820	CAGCTCCTCATCTGGATCCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	879
Qy	601	ATCTTGAATCTCAATATTGTGTACAAGCTCAGAGGAGAGCAATTTTGTCTCCGTGGC	660
Db	880	ATCTTGAATCTCAATATTGTGTACAAGCTCAGAGGAGAGCAATTTTGTCTCCGTGGC	939
Qy	661	TACTTCCACGGGAGAACCCGACCTCTGTTGTTTACCATTTACCTCCATCTTTGCGCACACTT	720
Db	940	TACTTCCACGGGAGAACCCGACCTCTGTTGTTTACCATTTACCTCCATCTTTGCGCACACTT	999
Qy	721	TGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGCGGCGCCCATCCAGAACCCGC	780
Db	1000	TGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGCGGCGCCCATCCAGAACCCGC	1059
Qy	781	TGGCTGGTACACATCATGTCGACATTTGCCAATGCTAGCTAGCCCTTCTGAAACAGCCATC	840
Db	1060	TGGCTGGTACACATCATGTCGACATTTGCCAATGCTAGCTAGCCCTTCTGAAACAGCCATC	1119
Qy	841	AACTTCTTCTCTTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGAGCGCCGACGCTC	900
Db	1120	AACTTCTTCTCTTACTGCTTTCATCAGCAAGCGGTTCCGACCATGCGAGCGCCGACGCTC	1179
Qy	901	AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCATATACTTTTTC	960
Db	1180	AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCATATACTTTTTC	1239
Qy	961	ATAACAGTATAGCCCTGGATCTCGCGGCAAACTCAACATGATCAAGATGCTGGTGTAC	1020
Db	1240	ATAACAGTATAGCCCTGGATCTCGCGGCAAACTCAACATGATCAAGATGCTGGTGTAC	1299
Qy	1021	CAGTATGACAAAATGGAACCTTATAAA	1050
Db	1300	CAGTATGACAAAATGGAACCTTATAAA	1329

VERSION AX664703.1 GI:29164463  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Lei, K.R., Kapeller-Libermann, R. and Glucksmann, M.  
TITLE of human proteins and uses thereof  
JOURNAL Patent: WO 02074960-A 12 26-SEP-2002;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..1062  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 93.0%; Score 1040.4; DB 6; Length 1062;  
Best Local Similarity 99.4%; Pred. No. 1.7e-249;  
Matches 1044; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ATGGAGCACAGCGCCCACTCGCAGCAACAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
Db 1 ATGGAGCACAGCGCCCACTCGCAGCAACAGCTCGCTCTCTTGGTGGTCCCGCGC 60  
QY 61 TCGGCTCGGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTCTTGCAGTCCGGT 120  
Db 61 TCGGCTCGGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTCTTGCAGTCCGGT 120  
QY 121 TTACACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGAGACAGAG 180  
Db 121 TTACACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGAGACAGAG 180  
QY 181 TCCTCCTACAATATCTCTTGGCACTCGTGTGCGCAATCTTGGTCTCTTTTCATA 240  
Db 181 TCCTCCTACAATATCTCTTGGCACTCGTGTGCGCAATCTTGGTCTCTTTTCATA 240  
QY 241 GTGTTGTGGATCTCTGTGGAAGATTCATCTTGAACATCAGATGCTCAGTCCGC 300  
Db 241 GTGTTGTGGATCTCTGTGGAAGATTCATCTTGAACATCAGATGCTCAGTCCGC 300  
QY 301 GACAGATCATAGAAGTGTGGAATTCATCATCCACCTCCATATGATGATCTGTA 360  
Db 301 GACAGATCATAGAAGTGTGGAATTCATCATCCACCTCCATATGATGATCTGTA 360  
QY 361 CGGTTAACCATTGACAGGTATATCGTGTCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
Db 361 CGGTTAACCATTGACAGGTATATCGTGTCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
QY 421 TACCAGCGCGCACCGGAAAGTCAATGTAAGTGTTCATCAGCTGCTTCTGACGAGC 480  
Db 421 TACCAGCGCGCACCGGAAAGTCAATGTAAGTGTTCATCAGCTGCTTCTGACGAGC 480  
QY 481 ATCCCTATTAATCTGTGGCGCAACATCTGAGCTGAAGACTACATCAGACCTCTGTCAT 540  
Db 481 ATCCCTATTAATCTGTGGCGCAACATCTGAGCTGAAGACTACATCAGACCTCTGTCAT 540  
QY 541 CAGCTCCTCATCTGGATCCACTGCTTACCGCTACCTGGTGGCTGCTTCCATCTTCTTC 600  
Db 541 CAGCTCCTCATCTGGTCCACTGCTTACCGCTACCGTGTACCGTGGCTGCTTCCATCTTCTTC 600  
QY 601 ATCTGGAATCAATCTGTGTACAGCTCAGAGAGAGAGCAATTTTGGTCTCCGTGGC 660  
Db 601 ATCTGGAATCAATCTGTGTACAGCTCAGAGAGAGAGCAATTTTGGTCTCTGTGGC 660  
QY 661 TACTCCAGGGAGACACCGCACTCTGTTTACCATTTACCTCTTCCGACACTT 720  
Db 661 TACTCCAGGGAGACACCGCACTCTGTTTACCATTTACCTCTTCCGACACTT 720  
QY 721 TGGGGCCCCCGCATCATCATGATCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780

Db 721 TGGGGCCCCCGCATCATCATGATCTTTTACCACTCTATATGGGGCGCCCATCCAGAACCGC 780  
QY 781 TGGCTGGTACACATCATGTCCGACATTCGCCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
Db 781 TGGCTGGTACACATCATGTCCGACATTCGCCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
QY 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGCGAGCGCCACAGCTC 900  
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACATGCGAGCGCCACAGCTC 900  
QY 901 AAGGCTTCTTCAAGTGCAGAGCAACCTCTACAGTTCTACCAACATCAATCACTTTTCC 960  
Db 901 AAGGCTTCTTCAAGTGCAGAGCAACCTCTACAGTTCTACCAACATCAATCACTTTTCC 960  
QY 961 ATACCAAGTGTAGCCCTCGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTATC 1020  
Db 961 ATACCAAGTGTAGCCCTCGATCTCGCGGCAAACTCACACTGATCAAGATGCTGTGTATC 1020  
QY 1021 CAGTATGACAAAATGGAACCTTATAAAA 1050  
Db 1021 CAGTATGACAAAATGGAACCTTATAAAA 1050  
RESULT 13  
AX664701  
LOCUS 1526 bp DNA linear PAT 22-MAR-2003  
DEFINITION Sequence 10 from Patent WO02074960.  
ACCESSION AX664701  
VERSION AX664701.1 GI:29164461  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
REFERENCE  
AUTHORS Lei, K.R., Kapeller-Libermann, R. and Glucksmann, M.  
TITLE 3850, 28472, 5495, 65507, 81588 and 14354 methods and compositions of human proteins and uses thereof  
JOURNAL Patent: WO 02074960-A 10 26-SEP-2002;  
Millennium Pharmaceuticals, Inc. (US)  
FEATURES  
Location/Qualifiers  
1..1526  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
139..1200  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAD80046.1"  
/translation="MEHTAHLAANSISLWSPGSACGLGFVPPVYVYLLQCLGLPAN  
ILTVILISQVARRQKSSYNLLAALADILVLFVDFVDFELLEDFILMOMQVPQDK  
IIEVLEFSSIHSTIWIIVPLDIDRYAVCHPKYHTVSYPARTKRVIVSVITCLPTS  
IPYWNFINWEDISTSVHVLVHCFTYQVPCSFIFLINSIVVYVLRKRSFRL  
CGYSTGKTALITFTSIFAILWAPIIMILYHLYGAPIQNRWLHVMISDIANMLAL  
NTANFLYCFISKRTMAATLKAPEKQKQPVQVFTNHNFSITSSPWISPHSHC  
IKMLVYQYDKNGKPIKVPSP"  
CDS  
Query Match 93.0%; Score 1040.4; DB 6; Length 1526;  
Best Local Similarity 99.4%; Pred. No. 1.7e-249;  
Matches 1044; Conservative 0; Mismatches 6; Indels 0; Gaps 0;  
QY 1 ATGGAGCACAGCGCCCACTCGCAGCAACAGCTCGCTGTCTTGGTGGTCCCGCGC 60  
Db 139 ATGGAGCACAGCGCCCACTCGCAGCAACAGCTCGCTGTCTTGGTGGTCCCGCGC 198  
QY 61 TCGGCTCGGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTCTTGCAGTCCGGT 120  
Db 199 TCGGCTCGGGCTTGGGTTTCGTGCGCGGTCTACTACAGCTCTTGCAGTCCGGT 258  
QY 121 TTACAGCAAAATATCTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGAG 180

Db	259	TTACGAGCAATATCTTGACAGTGAATCATCTCTCCAGCTGGTGGCGCAAGACAGAG	318
Qy	181	TCCTCTCAACAACTATCTCTTGCGCACTCGCTGCTGCCGACATCTTTGGTCCCTCTTTTTCATA	240
Db	319	TCCTCTTAGAACATATCTCTTGCGCACTCGCTGCTGCCGACATCTTTGGTCCCTCTTTTTCATA	378
Qy	241	GTGTTTGTGGACTTCCTGTTGGGAAGATTTCATCTTGAACATGCAAGATGCTTCAGGTCGCC	300
Db	379	GTGTTTGTGGACTTCCTGTTGGGAAGATTTCATCTTGAACATGCAAGATGCTTCAGGTCGCC	438
Qy	301	GACAAGATCATPAGAAGTGTGGAATTCTCATCTCCATCCACACCTCCATATGGATTACTGTGA	360
Db	439	GACAAGATCATPAGAAGTGTGGAATTCTCATCTCCATCCACACCTCCATATGGATTACTGTGA	498
Qy	361	CCGTTAACCAATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTATACCAACGCTCTCA	420
Db	499	CCGTTAACCAATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTATACCAACGCTCTCA	558
Qy	421	TACCCAGCCGCGCACCGGAAAGTCAATTGTAAGTGTTTATCATCACTGCTTCCTGACCAAGC	480
Db	559	TACCCAGCCGCGCACCGGAAAGTCAATTGTAAGTGTTTATCATCACTGCTTCCTGACCAAGC	618
Qy	481	ATCCCTTATTACTGGTGGCCCAACATCTGGACTGAAGACTATCATCAGCACCTCTGTGCAT	540
Db	619	ATCCCTTATTACTGGTGGCCCAACATCTGGACTGAAGACTATCATCAGCACCTCTGTGCAT	678
Qy	541	CACGTCTCATCTCGATCCACTGTCTCAACCGTCTACTGTGGTCCCTGCTCCATCTCTCTC	600
Db	679	CACGTCTCATCTGGTCCACTGTCTCAACCGTCTACAGGTGCCCTGCTCATCTCTCTTC	738
Qy	601	ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAAGCAATTTTCTGCTCCGTGGC	660
Db	739	ATCTTGAACTCAATCGTTGTGTACAAGCTCAGGAGGAAGCAATTTTCTGCTCTGTGGC	798
Qy	661	TACTTCAGGGGAAGACACCGCCATCTTGTGTACCAATTACCTCCATCTTTGGCACACATT	720
Db	799	TACTTCAGGGGAAGACACCGCCATCTTGTGTACCAATTACCTCCATCTTTGGCACACATT	858
Qy	721	TGGSCCCCGGCATCATCATGATTCTTTTACCACCTCTATGGGGCGCCCATCCAGAACCGC	780
Db	859	TGGSCCCCGGCATCATCATGATTCTTTTACCACCTCTATGGGGCGCCCATCCAGAACCGC	918
Qy	781	TGGCTGGTACACATCATGTCCGACATTTGCCAACATGCTAGGCCCTTCTGAAACACAGCCATC	840
Db	919	TGGCTGGTACACATCATGTCCGACATTTGCCAACATGCTAGGCCCTTCTGAAACACAGCCATC	978
Qy	841	AACTTCTTCTCTACTGTTTCATCAGCAAGCGGTTCCGCAACATATGGAGCGGCCACCGCTC	900
Db	979	AACTTCTTCTCTACTGTTTCATCAGCAAGCGGTTCCGCAACATATGGAGCGGCCACCGCTC	1038
Qy	901	AAGCTTTCTTCTCAAGTCCGAGACCACTGTACAGTTCTTACCAATCATTAAGTTTTC	960
Db	1039	AAGCTTTCTTCTCAAGTCCGAGACCACTGTACAGTTCTTACCAATCATTAAGTTTTC	1098
Qy	961	ATAACAGTAGCCCTCGATCTCCCGCGCAAACTCACACTGTCATCAAGATGCTGGTGTAC	1020
Db	1099	ATAACAGTAGCCCTCGATCTCCCGCGCAAACTCACACTGTCATCAAGATGCTGGTGTAC	1158
Qy	1021	CAGTATGACAAAATGGAAACCTTATAAA	1050
Db	1159	CAGTATGACAAAATGGAAACCTTATAAA	1188

[illegible]

QY 841 AACTTCTCTCTACTGCTTTTCATCAGCAAGCGGTTCCGCAACATGGCAGCCGCGCTC 900  
Db |||||  
841 AACTTCTCTCTACTGCTTTTCATCAGCAAGCGGTTCCGCAACATGGCAGCCGCGCTC 900  
QY 901 AAGGCTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCAATCACTTTTCC 960  
Db |||||  
901 AAGGCTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCAATCACTTTTCC 960  
QY 961 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACATGTCATCAAGATGCTGGTGAC 1020  
Db |||||  
961 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACATGTCATCAAGATGCTGGTGAC 1020  
QY 1021 CAGTATGACAAAATGGAACCTATATAA 1050  
Db |||||  
1021 CAGTATGACAAAATGGAACCTATATAA 1050

RESULT 15  
CQ737667  
LOCUS CQ737667 957 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 23601 from Patent WO02068579.  
ACCESSION CQ737667  
VERSION CQ737667.1 GI:42336565  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1  
Venter,C.J., Adams,M.C., Li,P.W. and Myers,B.W.  
Kits, such as nucleic acid arrays, comprising a majority of  
humanexons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 23601 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source  
1..957  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 83.1%; Score 929.8; DB 6; Length 957;  
Best Local Similarity 99.8%; Pred. No. 8.3e-222;  
Matches 931; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 118 GGTTTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAG 177  
Db |||||  
13 GGTTTTGACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAG 72  
QY 178 AAGTCTCTCTACAACTATCTCTTGGCACTCGTGTGGCGACATCTTGGTCTCTTTTC 237  
Db |||||  
73 AAGTCTCTCTACAACTATCTCTTGGCACTCGTGTGGCGACATCTTGGTCTCTTTTC 132  
QY 238 ATAGTGTGTTGGACTTCCTGTTGGAAGATTTTCATCTTGAAATGCAAGATGCTCAGGTC 297  
Db |||||  
133 ATAGTGTGTTGGACTTCCTGTTGGAAGATTTTCATCTTGAAATGCAAGATGCTCAGGTC 192  
QY 298 CCCGACAAGATCATAGAAGTGTGGAATTTCTCATCCATCCACACCTCCATATGGATTACT 357  
Db |||||  
193 CCCGACAAGATCATAGAAGTGTGGAATTTCTCATCCATCCACACCTCCATATGGATTACT 252  
QY 358 GTACCGTTAAACATTGACAGGTATATCGCTGTCTGCGCCCGCTCAAGTACCACAGGTC 417  
Db |||||  
253 GTACCGTTAAACATTGACAGGTATATCGCTGTCTGCGCCCGCTCAAGTACCACAGGTC 312  
QY 418 TCATACCCAGCCGCGACCCGGAAGTCAATTGTAAGTGTTCATCACCTGCTCTCTGACC 477  
Db |||||  
313 TCATACCCAGCCGCGACCCGGAAGTCAATTGTAAGTGTTCATCACCTGCTCTCTGACC 372  
QY 478 AGCATCCCCCTATTACTGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTG 537  
Db |||||  
373 AGCATCCCCCTATTACTGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTG 432

Search completed: February 12, 2005, 12:29:53  
Job time : 4817 secs

QY 538 CATCAGTCCCTCATCTGGATCACCTGCTTTACCGTCTACCTGGTGCCTGCTCCATCTTC 597  
Db |||||  
433 CATCAGTCCCTCATCTGGATCOACTGCTTTACCGTCTACCTGGTGCCTGCTCCATCTTC 492  
QY 598 TTCACTTTGAAACTCAATCAATTTGTGTACAAAGCTCAGGAGGAAAGCAATTTTGTCTCCGT 657  
Db |||||  
493 TTCACTTTGAAACTCAATCAATTTGTGTGTACAAAGCTCAGGAGGAAAGCAATTTTGTCTCCGT 552  
QY 658 GGCTACTCCAGGGGGAAGACCAACCGCATCTTTGTGTACCAATTACCTCCATCTTTGCCACA 717  
Db |||||  
553 GGCTACTCCAGGGGGAAGACCAACCGCATCTTTGTGTGTACCAATTACCTCCATCTTTGCCACA 612  
QY 718 CTTTGGGCCCCCGGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAAC 777  
Db |||||  
613 CTTTGGGCCCCCGGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAAC 672  
QY 778 CGCTGGCTGGTACACATCATGTCCGACATTTGCCAAACATGCTAGCCCTTTCTGAACAACAGCC 837  
Db |||||  
673 CGCTGGCTGGTACACATCATGTCCGACATTTGCCAAACATGCTAGCCCTTTCTGAACAACAGCC 732  
QY 838 ATCAAACTTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACATGGCAGCCGCGCAG 897  
Db |||||  
733 ATCAAACTTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCAACATGGCAGCCGCGCAG 792  
QY 898 CTCAGGCTTTCTTCAAGTGCAGAGCAACCTGTGTACAGTTTCTACCACAATCAATACTTT 957  
Db |||||  
793 CTCAGGCTTTCTTCAAGTGCAGAGCAACCTGTGTACAGTTTCTACCACAATCAATACTTT 852  
QY 958 TCCATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGTCATCAAGATGCTGGTG 1017  
Db |||||  
853 TCCATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGTCATCAAGATGCTGGTG 912  
QY 1018 TACCAAGTATGACAAAATGGAACCTATATAA 1050  
Db |||||  
913 TACCAAGTATGACAAAATGGAACCTATATAA 945

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 12, 2005, 12:29:57 ; Search time 3089 Seconds  
(without alignments)  
4583.978 Million cell updates/sec

Title: US-10-689-832-20

Perfect score: 1967

Sequence: 1 MHTAHLAANSLSSWSPG.....NDSKSSYOFDAIGACVIL 372

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp  
-O=/cgn2.1/USPTO.spool.p/US10689832/runat.08022005.112210.23251/app\_query.fasta\_1.519  
-DB=EST\_QPFT=fastap -SUFFIX=rest -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45  
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10689832 @CNC 1.1 5180 @runat.08022005.112210.23251 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*

1: gb\_est1:\*

2: gb\_est2:\*

3: gb\_hic:\*

4: gb\_est3:\*

5: gb\_est4:\*

6: gb\_est5:\*

7: gb\_est6:\*

8: gb\_gest1:\*

9: gb\_gest2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Match	Length	ID	Description
C 1	910	46.3	803	9	CC905917 t025011ba
2	606	30.8	806	9	AL178069 Tetraodon
3	509	25.9	824	9	CL137737 ISB1-110A
C 4	491	25.0	362	8	AZ513631 IM0359107
5	449.5	22.9	947	9	CL137737 ISB1-110A
6	444	22.6	867	9	CL137848 Tetraodon
7	433.5	22.0	813	8	BZ169352 CH230-259
8	426	21.7	898	9	AL93966 Tetraodon
9	414.5	21.1	515	1	AL921815 AL921815

C 10	397	20.2	448	8	AQ225693	AQ225693 HS 2009 B
11	381.5	19.4	1269	9	CL649540	CL649540 CH213-237
C 12	344	17.5	708	5	BUG10327	BUG10327 UT-M-DJ2-
13	314	16.0	588	8	AQ344035	AQ344035 RPC111-12
C 14	309	15.7	603	8	AQ351433	AQ351433 RPC111-11
15	249	12.7	1107	9	CNS0417P	AL295764 Tetraodon
16	241	12.3	1197	9	AY398777	AY398777 Mus muscu
17	239.5	12.2	1197	9	AY398776	AY398776 Pan trogl
18	238.5	12.1	1197	9	AY398775	AY398775 Homo sapi
19	231.5	11.8	1086	9	AY399291	AY399291 Homo sapi
20	231.5	11.8	2583	3	CR609171	CR609171 full-leng
21	218	11.1	1146	9	AY414479	AY414479 Mus muscu
22	218	11.1	2785	3	AK028281	AK028281 Mus muscu
23	217	11.0	2919	3	AK038620	AK038620 Mus muscu
24	217	11.0	2959	3	AK079529	AK079529 Mus muscu
25	216.5	11.0	2432	3	AK051189	AK051189 Mus muscu
26	216.5	11.0	2708	3	AK076275	AK076275 Mus muscu
27	215.5	11.0	1086	9	AY399293	AY399293 Mus muscu
28	215.5	11.0	2118	3	AK031109	AK031109 Mus muscu
29	215.5	11.0	2292	3	AK036597	AK036597 Mus muscu
30	215.5	11.0	2432	3	AK036690	AK036690 Mus muscu
31	213	10.8	863	5	BP161654	BP161654 BP161654
32	213	10.8	1511	3	CR601301	CR601301 full-leng
33	213	10.8	1603	3	CR614199	CR614199 full-leng
34	213	10.8	1612	3	CR594588	CR594588 full-leng
35	213	10.8	1619	3	CR610268	CR610268 full-leng
36	213	10.8	1621	3	CR614594	CR614594 full-leng
37	213	10.8	1626	3	CR598713	CR598713 full-leng
38	213	10.8	1629	3	CR596547	CR596547 full-leng
39	213	10.8	1629	3	CR623838	CR623838 full-leng
40	213	10.8	1640	3	CR619476	CR619476 full-leng
41	213	10.8	1642	3	CR594428	CR594428 full-leng
42	213	10.8	1650	3	CR605131	CR605131 full-leng
43	213	10.8	1682	3	CR614663	CR614663 full-leng
44	207.5	10.5	1062	9	AY400424	AY400424 Mus muscu
45	207.5	10.5	1200	9	AY416064	AY416064 Homo sapi

#### ALIGNMENTS

RESULT 1

CC905917/c

LOCUS

DEFINITION

ACCESION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

803 bp DNA linear GSS 08-AUG-2003  
t025011ba.r1 TAMBT Bos taurus genomic clone t025011ba, genomic  
survey sequence.

CC905917

CC905917.1 GI:33524850

GSS.

Bos taurus (cow)

Bos taurus

Bos taurus

Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

Bovinae; Bos.

1 (bases 1 to 803)

Lin, S., Najjar, F. Z., Adelson, D., Gill, C. A. and Roe, B. A.

Bovine BAC End Sequences from Library TAMBT

Unpublished (2003)

Contact: Bruce A. Roe

Advanced Center for Genome Technology

University of Oklahoma Department of Chemistry and Biochemistry

620 Parrington Oval, Room 208, Norman, OK 73019, USA

Tel: 405 325 4912

Fax: 405 325 7762

Email: broe@ou.edu

Class: BAC ends

High quality sequence start: 39

High quality sequence stop: 551.

Location/Qualifiers

1. 803

/organism="Bos taurus"

/mol\_type="genomic DNA"

/strain="Angus bull T A M U Shoshone Y6 11519666"

/db\_xref="taxon:9913"

/clone="t025o11ba"		Tetradontoidea; Tetraodontidae; Tetraodon.	
/sex="Male"		1	
/cell_type="Blood"		Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,	
/clone_libs="TAMBT"		Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,	
/note="Vector: pBelOBAC11; Site 1: HindIII; Site 2:		Saurin,W. and Weissenbach,J.	
HindIII; TAMBT Bovine BAC library (Male) produced by Texas		Estimate of human gene number provided by genome-wide analysis	
A&M University, Department of Animal Science."		using Tetraodon nigroviridis DNA sequence	
ORIGIN		Nat. Genet. 25 (2), 235-238 (2000)	
Alignment Scores:		20296633	
Pred. No.:	1.52e-86	Length:	803
Score:	910.00	Matches:	173
Percent Similarity:	92.63%	Conservative:	3
Best Local Similarity:	91.05%	Mismatches:	13
Query Match:	46.26%	Indels:	1
DB:	9	Gaps:	0
US-10-689-832-20 (1-372) x CC905917 (1-803)		20359837	
QY 162 ProTyTyTrrpTrpProAsnIleTrrpThrGluAspTyrlleSerThr-SerValHisH1 181		10899143	
Db 765 CCTAACAGCTGGGGCCCCAACCCCTGGACTTGAGATTACATCAGCACGGTCCATGCA 706		3 (bases 1 to 806)	
QY 181 sValLeuIleTrrpIleHisCysPheThrValTyrlleuValProCysSerIlePhePhe 201		Genoscope.	
Db 705 TGTCTTATCTGAATCTACTGCTCCCTGTGTACTGGGTACCTGTTCCATATCTTCAT 646		Direct Submission	
QY 201 eLeuAsnSerIleIleValTyrlsLeuArgGlyssSerAsnPheArgLeuArgGlyTy 221		Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :	
Db 645 CTGAAACTCCATCATTTGTGTACAGCTCAGGAGGAAGCAATTTTGCCTCCCTGGCTA 586		BP 191 91006 EVRY cedex - FRANCE (E-mail : secref@genoscope.cns.fr	
QY 221 rSerThrGlylysrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTr 241		- Web : www.genoscope.cns.fr)	
Db 585 CTCCACGGGGAAAACCCGCTGCATCTTGTTTACCACCACTCCATCTTTGCCACCCCTG 526		This sequence is a single read and was generated as part of a large	
QY 241 pAlaProArgIleIleMetIleLeuTyrlsLeuTyrlsGlyAlaProIleGlnAsnArgTr 261		scale clone-end sequencing project of the Tetraodon nigroviridis	
Db 525 GGCCCCACGATCATCATGATCTCTACCACTTTACGGGAGCGCCATCCAGAACCGCTG 466		genome. For more information, please take a look at	
QY 261 pLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAs 281		http://www.genoscope.cns.fr/Tetraodon.	
Db 465 GCTGGTGCACGTGATGCGGATGTTGGCAACATGCTGGCCCTTCTGAAACAGGCCATCA 406		Location/Qualifiers	
QY 281 nphePheLeuTyrlsCysPheIleSerIlyeArgPheArgThrMetAlaAlaThrLeuLy 301		1. .806	
Db 405 CTTCCTCTCTATTGCTTCATCAGCAACGCGTTCGTCGCCATGGCAGCGCTACTCTCA 346		/organism="Tetraodon nigroviridis"	
QY 301 sAlaPhePheLyssGlnLysGlnProValGlnPheTyrlsThrAsnHisAsnPheSerI 321		/mol_type="genomic DNA"	
Db 345 GGCCTTCTTCAAGTCCAGAGCAACCCGTTCACTTCTACCAACCACTATCTTTCCAT 286		/db_xref="taxon:99883"	
QY 321 eThrSerSerProThrIleSerProAlaAsnSerHisCysIleLyssMetLeuValTyrl 341		/clones="227D19"	
Db 285 AACAGTAGCCCTGGATCTCACCGGCCAATCTACACTGTATCAAGATGCTGTGTACCA 226		/clone_lib="G"	
QY 341 nTyrlsPlyssAnGlyLyssProIleLys 350		/note="Genoscope sequence ID : COAG227CB10LP1-end : T7"	
Db 225 GTATGACAAAATGGAAAAACCTATAAA 198		Alignment Scores:	
RESULT 2		Pred. No.:	5.43e-54
CNS022E4		Score:	606.00
LOCUS		Percent Similarity:	65.99%
DEFINITION		Best Local Similarity:	44.94%
Tetraodon nigroviridis genome survey sequence T7 end of clone		Query Match:	30.81%
227D19 of library G from Tetraodon nigroviridis, genomic survey		DB:	9
sequence.		US-10-689-832-20 (1-372) x CNS022E4 (1-806)	
ACCESSION		QY 57 ArgArgGlnlyssSerSerTyrlsAsnTyrlleuAlaLeuAlaAlaAspIleLeuVal 76	
VERSION		Db 5 AAGACCAAGAGGCGCTGTACTACTCTCGWGGCGTGACAGGCTCCGACATCTCTCC 64	
KEYWORDS		QY 77 LeuphePheIleValPheValAspPheIleuGluAspPheIleLeuAsnMetGlnMet 96	
SOURCE		Db 65 CAGCTCTTCATCATCTTCGTGGGCTTCTGTGTGAGACGCGCGGTTTCCACCGGAGGC 124	
ORGANISM		QY 97 ProGlnValProAspLyssIleIleGluValleuGluPheSerSerIleHisThrSerIle 116	
Tetraodon nigroviridis		Db 125 CCGCGCTCTCTTACACTCAGTCAGCGCGCGAGTTTGCCGCCAACACGCGCTCCATC 184	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		QY 117 TrpIleThrValProLeuThrIleAspArgTyrlleAlaValCysHisProLeuLysTyrl 136	
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		Db 185 TGGTCCACCGTCCCTCCCTCACCGTGACCGCTACGTGGCGGTGTGTGCCACCCCTCTCCAC 244	
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;		QY 137 HisThrValSerTyrlsProAlaArgThrArgLyssValIleValSerValTyrlleThrCys 156	
		Db 245 AGGCAGATCAGTACCGCGCCCGACACGAGGAGTATCGCGGTGGTCTCTGGTGTCTGTCG 304	
		QY 157 PheLeuThrSerIleProTyrlsTyrlsTrpTrpProAsnIleTrpThrGluAspTyrlleSer 176	
		Db 305 CTGGCCTCGGCGCTCTCTCTGTGTGTGTCCGACATGTGTGAGGAACACGCCACCGCGC 364	





JOURNAL COMMENT

plasmid inserts  
Unpublished (2000)  
Contact: Robert B. Weiss  
University of Utah Genome Center  
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT  
84112, USA  
Tel: 801 585 5606  
Fax: 801 585 7177  
Email: gdunn@genetics.utah.edu  
Insert Length: 10000 Std Error: 0.00  
Plate: 0359 row: L column: 07  
Seq primer: CACACGGAACACGATGACC  
Class: plasmid ends  
High quality sequence stop: 362.  
Location/Qualifiers  
1. 362  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0359L07"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

FEATURES  
source  
1. 362  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/strain="C57BL/6J"  
/db\_xref="taxon:10090"  
/clone="UUCG1M0359L07"  
/sex="Male"  
/lab\_host="E. Coli strain XL10-Gold, T1-resistant, F-"  
/clone\_lib="Mouse 10kb plasmid UUCG1M library"  
/notes="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource  
(http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (gi|4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid RL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

ORIGIN

Alignment Scores:  
Pred. No.: 3.48e-42 Length: 362  
Score: 491.00 Matches: 91  
Percent Similarity: 95.88% Conservative: 2  
Best Local Similarity: 93.81% Mismatches: 4  
Query Match: 24.96% Indels: 0  
DB: 8 Gaps: 0

US-10-689-832-20 (1-372) x AZ5113631 (1-362)

Qy 254 GlyAlaProIleGlnAsnArgTrpLeuValHisIleMetSerAspIleAlaAsnMetLeu 273  
Db 361 GGAGCACCCTCCAGAACCCCTGGCTGGTGCACATCATGTGGATGTGCCAACATGCTA 302  
Qy 274 AlaLeuLeuAnThrAlaIleAsnPhePheLeuTyrCysPheIleSerLysArgPheArg 293  
Db 301 GCCCTTNTGAACACAGCATCAACTCTTCTCTACTGCTTCATCAGCAAGCGCTTCGT 242  
Qy 294 ThrMetAlaAlaAlaThrLeuLysAlaPhePheLysCysGlnLysGlnProValGlnPhe 313  
Db 241 ACCATGGCAGCTGCCACACTCAAGCCCTGTTCAGTGTCAAGACAGCGCTGTACAGTTC 182  
Qy 314 TyrThrAsnHisAsnPheSerIleThrSerSerProTyrPheSerProAlaAsnSerHis 333  
Db 181 TATACCAACCATTAACATTTTCCATACAAAGTAGTCCCTGGATCTCACCAAGCAACTCACAC 122  
Qy 334 CysIleLysMetLeuValTyrGlnTyrAspLysAsnGlyLysProIleLys 350

Db 121 TGCATCAAGATGCTGGTGTACCATGATGACAAACATGGAAGCCTATAAAA 71

RESULT 5  
CNS02BVQ  
LOCUS  
DEFINITION  
CNS02BVQ 947 bp DNA linear GSS 01-SEP-2000  
Tetraodon nigroviridis genome survey sequence PUC-ORI end of clone  
254A20 of library G from Tetraodon nigroviridis, genomic survey  
sequence.  
AL190367  
GI:7828471  
GSS: genome survey sequence.  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Tetraodon nigroviridis  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
Tetraodontoidea; Tetraodontidae; Tetraodon.  
1  
Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,  
Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,  
Saurin,W. and Weissenbach,J.  
Estimate of human gene number provided by genome-wide analysis  
using Tetraodon nigroviridis DNA sequence  
Nat. Genet. 25 (2), 235-238 (2000)  
20296633  
10835645  
2  
Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,  
Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,  
Saurin,W., Bernot,A. and Weissenbach,J.  
Characterization and repeat analysis of the compact genome of the  
freshwater pufferfish Tetraodon nigroviridis  
Genome Res. 10 (7), 939-949 (2000)  
20359837  
10899143  
3 (bases 1 to 947)  
Genoscope.  
Direct Submission  
Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
This sequence is a single read and was generated as part of a large  
scale clone-end sequencing project of the Tetraodon nigroviridis  
genome. For more information, please take a look at  
http://www.genoscope.cns.fr/tetraodon.  
Location/Qualifiers  
1. 947  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="254A20"  
/clone\_lib="G"  
/note="Genoscope sequence ID : COAG254BA10SP1-end :  
PUC-ORI"

FEATURES  
source  
1. 947  
/organism="Tetraodon nigroviridis"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:99883"  
/clone="254A20"  
/clone\_lib="G"  
/note="Genoscope sequence ID : COAG254BA10SP1-end :  
PUC-ORI"

ORIGIN

Alignment Scores:  
Pred. No.: 3.92e-37 Length: 947  
Score: 449.50 Matches: 85  
Percent Similarity: 79.17% Conservative: 10  
Best Local Similarity: 70.83% Mismatches: 22  
Query Match: 22.85% Indels: 3  
DB: 9 Gaps: 1

US-10-689-832-20 (1-372) x CNS02BVQ (1-947)

Qy 233 IleThrSerIlePheAlaThrLeuTrpAlaProGlieIleMetIleLeuTyrHisLeu 252  
Db 12 ATCACCTCCATTTTGGCGTTCTGTGGCGCCGCCACCTCGATGATCTCTACCACTTC 71  
Qy 253 TyrGlyAlaProIleGlnAsnArgTrp-----LeuValHisIleMetSerAspIle 269

```

Db      72  TACTCCCTCCGCCACCTCGCGGGCGCGCGCTGCTGCACATTCCTCACCGATCTT 131
Qy      270 AlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhelLeuValPheValaspPheLeu 86
Db      132 GCCAACATGTGGCGTTGCTCAACACGGGTGTTAACTTCTCTCTACTGTTTCATCAGC 191
Qy      290 LysArgPheArgThrMetAlaAlaAlaThrLeuLysAlaPhePheLysCysGlnLysGln 309
Db      192 WAGCGTTTCGGGGCATGGCGGCCAACGCTGCTCGAGCGCTGCTCCACTGCAGGAAGCAG 251
Qy      310 ProValGlnPheThrThrAsnHisAsnPhelThrSerSerProThrPheLeu 329
Db      252 CTGACCGCTTCTACCCGACCAACAACTTTTCATCAGAGTGCCTCGGTCTCACCA 311
Qy      330 AlaAsnSerHisCysIleLysMetLeuValTyrglnTyrglnTyrglnTyrglnTyrgln 349
Db      312 GCAAATCCCACTGCATCAAGATGCTGTGTACCACTAGCAGACAAACCGGAAGCCCATC 371

RESULT 6
LOCUS   CL137848
DEFINITION  ISB1-110C15 T7.1 ISB1 Xenopus tropicalis genomic clone ISB1-110C15,
genomic survey sequence.
ACCESSION  CL137848
VERSION    CL137848.1 GI:40631483
KEYWORDS   GSS.
SOURCE     Xenopus tropicalis (western clawed frog)
ORGANISM   Xenopus tropicalis
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
           Xenopodinae; Xenopus; Silurana.
REFERENCE  1 (bases 1 to 867)
AUTHORS   Krenitzki, C., Carter, J., McPherson, J., Warren, W., Graves, T.,
           Mardis, E. and Wilson, R.
TITLE     A physical map of the xenopus tropicalis genome
JOURNAL   Unpublished (2003)
COMMENT   Contact: Richard K Wilson
           Genome Sequencing Center
           Washington University School of Medicine
           Email: submissions@watson.wustl.edu
           Insert Length: 75000 Std Error: 0.00
           Seq primer: T7 TAATACGACTCACTATAGG
           Class: BAC ends
           High quality sequence start: 11
           High quality sequence stop: 673.
FEATURES   Location/Qualifiers
           1..867
           /organism="Xenopus tropicalis"
           /mol_type="genomic DNA"
           /db_xref="taxon:8364"
           /clone="ISB1-110C15"
           /clone_lib="ISB1"
           /note="Vector: pBelOBAC11; ISB-1 Xenopus tropicalis BAC
           Library Segment 1"
ORIGIN
Alignment Scores:
Pred. No.: 1,34e-36 Length: 867
Score: 444.00 Matches: 92
Percent Similarity: 61.06% Conservative: 46
Best Local Similarity: 40.71% Mismatches: 62
Query Match: 22.57% Indels: 26
DB: 9 Gaps: 4

US-10-689-832-20 (1-372) x CL137848 (1-867)
Qy      27 PheValProValValTyrglnSerLeuLeuLeuCysLeuGlyLeuProAlaAsnLeuLeu 46
Db      172 TTCTCTCTACAGTG-----AACATGCTG 195
Qy      47 ThrValIleLeuSerGlnLeuValAlaArgArgGlnLysSerSerTyrglnTyrgln 66
Db      196 ACTGCTGTGGCACTCTCAAGGCTGGCATCGAACAACAAAGTCTCTCTACAGTTACCTG 255

```

```

Qy      67 LeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIleValPheValaspPheLeu 86
Db      256 TTGGCTCTTACCATCTCAGACATCTTGAGCCAGATTTTCATCATTTTGTGGCTTATC 315
Qy      87 LeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAspLysIleleleGluVal 106
Db      316 CTGCAACAGCAATACTCCACGCAAGGTGCCCTTAATCCATCGGTGAGTGC 375
Qy      107 LeuGluPheSerSerIleHisThrSerIleTrrPheThrValProleuThrIleAspArg 126
Db      376 CTGAGTTCTCTCTTAATCATGCATCTATTGGGTGACAGTGATCTGACCGTGGACCGT 435
Qy      127 TyrIleAlaValCysHisProLeuLysTyrglnHisThrValSerTyrglnAlaArgThrArg 146
Db      436 TATGTGGCATATGCTATCGCTACAATATCGCTCTTCTCTTACCAGAGCAACCGT 495
Qy      147 LysValIleValSerValTyrglnThrCysPheLeuThrSerIleProTyrglnTyrgln 166
Db      496 AGAGTCATTTGCTTCTCTCTCATGCTTTGTGACCGGTGATCCCTTCTACTGTGG 555
Qy      167 ProAlaIleTrrPheThrGluAspTyrglnSerThrSerValHisValleuIleTrrPhe 186
Db      556 AGTGATGTTGGAGGATCCGCTATCCCGAGGATGTGACCTTATCTCAAGTGGACC 615
Qy      187 HisCysPheThrValTyrglnLeu--ValProCysSerIlePhePheIleLeuAsnSerIleI 206
Db      616 CACTGCTTTATTAATATATCTTCCATTCCTCGACCATATTTTTCATTAACCAACTCTG 675
Qy      206 leValTyrgln-----LysLeuArgArgLysSer-----Asp 216
Db      676 GTCAATTATCTACCAAGGCTGAGGAAAAAAGCGGGTCCCAAGAGAGTTGCCCATCCT 735
Qy      216 heArgLeuArgGlyTyrglnSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerI 236
Db      736 TTCCTAAATGGGCCAAGAA-----CCACACGATCAATTTCTTCTTCTTTGGAGACAA 789
Qy      236 lePheAlaThrLeu 240
Db      790 TTTACCCCTCCCTG 803

RESULT 7
LOCUS   BZ169352
DEFINITION  BZ169352 TJ CHORI-230 Segment 2 Rattus norvegicus genomic clone
CH230-259E22, genomic survey sequence.
ACCESSION  BZ169352
VERSION    BZ169352.1 GI:23810547
KEYWORDS   GSS.
SOURCE     Rattus norvegicus (Norway rat)
ORGANISM   Rattus norvegicus
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
           Rattus.
REFERENCE  1 (bases 1 to 813)
AUTHORS   Zhao, S., Shetty, J., Shatsman, S., Tseng, G., Geer, K.,
           Shvartsbeyn, A., Gebregorgis, E., Overton, L., Russell, D., Chen, D.,
           Riggs, P., de Jong, P. and Fraser, C.M.
           Rat BAC End Sequences from Library CHORI-230 MbOI segment
           Unpublished (1999)
           Other GSSs: CH230-259E22.TV
           Contact: Shaying Zhao
           Department of Eukaryotic Genomics
           The Institute for Genomic Research
           9712 Medical Center Dr., Rockville, MD 20850, USA
           Tel: 301 838 0200
           Fax: 301 838 0208
           Email: szhao@tigr.org
           Clones are derived from the rat BAC library CHORI-230
           (http://www.chori.org/bacpac/rat230.htm). For BAC library
           availability, please contact Fietser de Jong (pdejong@mail.cho.org).
           Clones may be purchased from BACPAC Resources
           (http://www.chori.org/bacpac/or ering_information.htm). BAC end

```

page: [http://www.tigr.org/tdb/bac\\_ends/rat/bac\\_end\\_intro.html](http://www.tigr.org/tdb/bac_ends/rat/bac_end_intro.html)  
 Plate: 259 row: E column: 22  
 Seq primer: SP6  
 Class: BAC ends.

## FEATURES

source  
 1. .813  
 Location/Qualifiers  
 /organism="Rattus norvegicus"  
 /mol\_type="genomic DNA"  
 /strain="BN/SSHsd/MCW"  
 /db\_xref="taxon:10116"  
 /clone="CH230-259E22"  
 /sex="Female"  
 /cell\_type="Brain"  
 /clone\_lib="CHORI-230 Segment 2"  
 /note="Vector: pTARBAC1.3; Site\_1: MboI; Site\_2: MboI;  
 CHORI-230 Rat (BN/SSHsd/MCW) BAC library produced by  
 Pieter de Jong"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.62e-35 Length: 813  
 Score: 433.50 Matches: 90  
 Percent Similarity: 88.57% Conservative: 3  
 Best Local Similarity: 85.71% Mismatches: 5  
 Query Match: 22.04% Indels: 7  
 DB: 8 Gaps: 1

US-10-689-832-20 (1-372) x BZ169352 (1-813)

QY 37 LeuCysLeuGly-----LeuProAlaAsnIleLeuThrValIle 49  
 DB 499 CTTTGCTTGGGATAACTTTGATTCTTCTCTCCACAGCAGATATCTTGACAGTCATC 558  
 QY 50 IleLeuSerGlnLeuValAlaArgGlnIlySerSerTyrAsnTyrLeuLeuAlaLeu 69  
 DB 559 ATCTCTCTCAACTGGTGGCAGAGACAGAAAGTCCTCTCAACTATCTTCTGCACTT 618  
 QY 70 AlaAlaAlaAspIleLeuValLeuPhePheIleValPheValAspPheLeuLeuGluAsp 89  
 DB 619 GCTGCTGCCGACATCTGGTCTCTTTTTCATCGCTTTTGTGGATTTCTTGTAGAAGAC 678  
 QY 90 PheIleLeuAsnMetGlnMetProGlnValProAlaPheIleIleGluValLeuGluPhe 109  
 DB 679 TTCAATTTTGACCATGACATGCTCCGATCCCTGACAGATCATAGAGTGTAGATTC 738  
 QY 110 SerSerIleHisThrSerIleTrpIleThrValProLeuThrIleAspArgTyrIleAla 129  
 DB 739 TCCTCCATCCACACTCTATTGTGATTACAGTCCCTTAAACAGTTGATAGGTATATCGCA 798  
 QY 130 ValCysHisProLeu 134  
 DB 799 GTCTGTCAACCCACTC 813

## RESULT 8

## CNS02ENP

## LOCUS

DEFINITION Tetraodon nigroviridis genome survey sequence, PUC-ori end of clone 262F08 of library G from Tetraodon nigroviridis, genomic survey sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodontidae; Tetraodon.

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Bouneau, L., Fisher, C., Bernot, A., Fizames, C., Wincker, P., Brottier, P., Quetier, F., Saurin, W. and Weissenbach, J.

Estimate of human gene number provided by genome-wide analysis

## TITLE

using Tetraodon nigroviridis DNA sequence  
 Nat. Genet. 25 (2), 235-238 (2000)

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

Roest Crolius, H., Jaillon, O., Dasilva, C., Ozouf-Costaz, C., Fizames, C., Fischer, C., Bouneau, L., Billault, A., Quetier, F., Saurin, W., Bernot, A. and Weissenbach, J.  
 Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

## JOURNAL

```

Db      792  GTCTCATCTGGACCCACCTGACCATCATCTCTCTGCTGCTGACGATCTTCTCGGTC 851
QY      202  LeuAsnSerIleIleValTyrLysLeuArg 211
Db      852  CTCACACTCTTTCATCATCCACAGCTGAGG 881

RESULT 9
LOCUS   AL921815
DEFINITION AL921815 PUR-21+22 Danio rerio cDNA clone 101-F03-2, mRNA sequence.
ACCESSION AL921815
VERSION   AL921815.1 GI:23188395
KEYWORDS EST.
SOURCE   Danio rerio (zebrafish)
ORGANISM Danio rerio
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
          Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 515)
AUTHORS   Lo,J., Lee,S., Xu,M., Liu,F., Ruan,H., Eun,A., He,Y., Ma,W.,
          Wang,W., Wen,Z., and Peng,J.
          15000 unique zebrafish EST clusters and their future use in
          microarray for profiling gene expression patterns during
          embryogenesis
JOURNAL   Genome Res. 13 (3), 455-466 (2003)
MEDLINE   22505427
PUBMED    12618376
COMMENT   Contact: Peng J
          Lab of Functional Genomics
          Institute of Molecular and Cell Biology
          30 Medical Drive, Singapore, 117609, Singapore
          Email: pengj@imcb.a-star.edu.sg
          Clone requests: info@openbiosystems.com
          Open Biosystems,
          6705 Odyssey Drive, Huntsville, AL 35806.
          Location/Qualifiers
FEATURES             source
     source           1..515
                     /organism="Danio rerio"
                     /mol_type="mRNA"
                     /strain="local wildtype"
                     /db_xref="taxon:7955"
                     /clone="101-F03-2"
                     /tissue_type="whole embryo or fish"
                     /dev_stage="mixed stages"
                     /clone_lib="PUR-21+22"

ORIGIN
Alignment Scores:
Pred. No.:          9.03e-34          Length:          515
Score:              414.50           Matches:         104
Percent Similarity: 37.46%           Conservative:    14
Best Local Similarity: 33.02%         Mismatches:     29
Query Match:        21.07%           Indels:         168
DB:                  1                Gaps:           4

US-10-689-832-20 (1-372) x AL921815 (1-515)

QY      1  MetGluHisThrHisAlaHisLeuAlaAAsnSerSerLeuSerTyrTrpSerProGly 20
Db      60  ATAGAGCACAGCCACATCTTCACCTGCTGACCACTCCAGCTCTCGGAGCCCT--- 116
QY      21  SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db      117  CGCGGTTCCTCTCGGACAGTTTCAGTCGTCCTACTACAGTCTCTTACTGTCCTCGGC 176
QY      41  LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAArgGlnLys 60
Db      177  CTGCAGCAAAACATCTCAGCGTGATCATCTTTCTCAGCTG----- 218
QY      61  SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAAspIleLeuValLeuPhePheIle 80
Db      218  ----- 218

```

```

QY      81  ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      218  ----- 218
QY      101  AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db      218  ----- 218
QY      121  ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db      218  ----- 218
QY      141  TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      218  ----- 218
QY      161  IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db      218  ----- 218
QY      181  HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db      218  ----- 218
QY      201  IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPhe-ArgLeuArgGln 220
Db      219  -----CTCCCGCTCCGAGG 233
QY      220  YTrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIle 240
Db      234  ATACTCGACCGGAAAGACACCGCTATCTTTTAGCATCACCTCAGTGTTCGCGCTTTT 293
QY      240  uTrpAlaProArgIleIleMetIleLeuTyrHisLeuTyr-----GlyAlaPr 256
Db      294  ATGGCGGCTCTACACTATGATTTCTTATCATCTTTACACGGTGCACACGCGATGCC 353
QY      256  oileGlnAsnArgTrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeu 276
Db      354  GGGTCGGCTAGA---CTCTGCAATTGGTAACGGATGTAGCGAACATGCTTGTCTACT 410
QY      276  uAsnThrAlaIleAsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAl 296
Db      411  CAACACCGGGGTCAACTTCTTCTCTACTGCTTTTATAAGCAAGCGCTTTTCGGAGGATGCG 470
QY      296  aAlaAlaThrLeuLysAlaPhePheLysCysGlnLysGlnPro 310
Db      471  CGGGACGGTGCTCAAGCCTTTTTCGCTGCAGAAACAGCCG 513

```

```

RESULT 10
LOCUS   AQ225693/c
DEFINITION HS_2009_B2_B08_F7 CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2009 Col=16 Row=D, genomic survey sequence.
ACCESSION AQ225693
VERSION   AQ225693.1 GI:3650922
KEYWORDS GSS
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
          1 (bases 1 to 448)
REFERENCE Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
          Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
          Hood,L.
          Sequence-tagged connectors: A sequence approach to mapping and
          scanning the human genome
          Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
JOURNAL 99380589
MEDLINE 1049764
PUBMED
COMMENT Contact: Mahairas GG, Wallace JC, Hood L
          High Throughput Sequencing Center

```



Db	123	TGGATCTCACCAGCAAACTCACA	CTGCATCAAGATGCTGGTGTACCAAGTATGACAACAT	102
QY	346	GlytysProfileLys	350	
Db	183	GGAAAGCCTATATAA	197	
RESULT 13				
AQ344035/c				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
COMMENT				
FEATURES				
source				
ORIGIN				
Alignment Scores:				
Pred. No.:				
Score:				
Percent Similarity:				
Best Local Similarity:				
Query Match:				
DB:				
US-10-689-832-20				
QY	43	AlaasnlleuthrValllelleLeuSerGlnleuValAlaArgGlnLysSerSer	62	
Db	199	GCAATATCTTGACAGTGATCATCTCTCCACGTGGTGGCAAGACAGAGTCTCTCC	140	
QY	63	TyrAsnTyrlleuLeuAlaLeuAlaAlaAspilleuValleuPheIleValPhe	82	
Db	139	TACAACTATCTCTTGGGACATCGCTGCTGCCGACATCTTGGTCTCTCTTTTCATAGTGT	80	
QY	83	ValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAspLys	102	

```

Db      79 GTGACTTCCTGTTGGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCTCCGACAAG 20
Qy      103 lleleGluValLeuGlu 108
Db      19 ATCATAGAAGTGTGGAA 2

RESULT 14
AQ351433/c
LOCUS   AQ351433
DEFINITION   AQ351433 603 bp DNA linear GSS 07-MAY-1999
              RPC111-113113.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-113113,
              genomic survey sequence.
ACCESSION   AQ351433
VERSION     AQ351433.1 GI:4178768
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 603)
AUTHORS     Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and
            Venter,J.C.
TITLE       Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready
            Map Building
JOURNAL     Unpublished (1997)
COMMENT     Other GSSs: RPC111-113113.TV
            Contact: Shaying Zhao, William Niernan, Mark Adams
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: hbe@tigr.org
            Clones are derived from the human BAC library RPCI-11. For BAC
            library availability, please contact Pieter de Jong
            (pieter@dejong.med.buffalo.edu). Clones may be purchased from
            BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from
            Research Genetics (info@resgen.com). BAC end search page:
            http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html
            Seq primer: SP6
            Class: BAC ends.

FEATURES             Location/Qualifiers
     source           1..603
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="GDB:7543212"
                     /db_xref="taxon:9606"
                     /clone="RPCI-11-113113"
                     /sex="Male"
                     /cell_type="Lymphocytes"
                     /clone_lib="RPCI-11"
                     /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                     RPC111 Human Male BAC Library"

ORIGIN
Alignment Scores:
Pred. No.:          2.25e-22      Length:      603
Score:              309.00        Matches:     65
Percent Similarity: 100.00%       Conservative: 0
Best Local Similarity: 100.00%    Mismatches:  0
Query Match:        15.71%        Indels:     0
DB:                  8            Gaps:       0

US-10-689-832-20 (1-372) x AQ351433 (1-603)

Qy      43 AlaAniLeuThrValIleLeuSerGlnLeuValAlaArgArgGlnLysSerSer 62
Db      195 GCATATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGACAGAGTCTCTCC 136

Qy      63 TyrAnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleValPhe 82
Db      135 TACAACATATCTCTTGCACTCGCTGCTGCCGACATCTTGGTCCCTCTTTTCATAGTGT 76

Qy      83 ValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAspLys 102

```

```

Db      75 GTGACTTCCTGTTGGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCTCCGACAAG 16
Qy      103 lleleGluValLeu 107
Db      15 ATCATAGAAGTGTCTG 1

RESULT 15
CNS04L7F
LOCUS   CNS04L7F
DEFINITION   CNS04L7F 1107 bp DNA linear GSS 01-SEP-2000
              Tetraodon nigroviridis genome survey sequence T7 end of clone
              118A07 of library G from Tetraodon nigroviridis, genomic survey
              sequence.
ACCESSION   AL295764
VERSION     AL295764.1 GI:8034344
KEYWORDS    GSS; genome survey sequence.
SOURCE      Tetraodon nigroviridis
ORGANISM    Tetraodon nigroviridis
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontoidea; Tetraodontidae; Tetraodon.
REFERENCE   1
AUTHORS     Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
            Bernot,A., Bizames,C., Wincker,P., Brottier,P., Quetier,F.,
            Saurin,W. and Weissenbach,J.
TITLE       Estimate of human gene number provided by genome-wide analysis
            using Tetraodon nigroviridis DNA sequence
JOURNAL     Nat. Genet. 25 (2), 235-238 (2000)
MEDLINE     20296633
PUBMED      10835645
REFERENCE   2
AUTHORS     Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
            Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
            Saurin,W., Bernot,A. and Weissenbach,J.
TITLE       Characterization and repeat analysis of the compact genome of the
            freshwater pufferfish Tetraodon nigroviridis
JOURNAL     Genome Res. 10 (7), 939-949 (2000)
MEDLINE     20359837
PUBMED      10899143
REFERENCE   3 (bases 1 to 1107)
AUTHORS     Genoscope.
TITLE       Direct Submission
JOURNAL     Submitted (12-APR-2000) Genoscope - Centre National de Sequencage :
            BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
            - Web : www.genoscope.cns.fr)
COMMENT     This sequence is a single read and was generated as part of a large
            scale clone-end sequencing project of the Tetraodon nigroviridis
            genome. For more information, please take a look at
            http://www.genoscope.cns.fr/tetraodon.
            Location/Qualifiers
            1..1107
            /organism="Tetraodon nigroviridis"
            /mol_type="genomic DNA"
            /db_xref="taxon:99883"
            /clone="118A07"
            /clone_lib="G"
            /note="Genoscope sequence ID : COBGL18AA04LP1-end : T7"

ORIGIN
Alignment Scores:
Pred. No.:          1.45e-15      Length:      1107
Score:              249.00        Matches:     41
Percent Similarity: 81.43%       Conservative: 16
Best Local Similarity: 58.57%    Mismatches:  13
Query Match:        12.66%        Indels:     0
DB:                  9            Gaps:       0

US-10-689-832-20 (1-372) x CNS04L7F (1-1107)

Qy      224 GlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTrpAlaPro 243
Db      71 GGGAGAGCAGCGCGCATGCTGCGCCATCACCTCCGTCCTCTCTGTGCTCTGGGACCC 130

```



```
Qy 244 ArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArgTrrpLeuVal 263
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Qy 131 AGGACGGTGGTGGTCACTACCACTGTACGTGTCGGTTCACCGCGACTGGCGGTC 190
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Qy 264 HisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePhe 283
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Qy 191 CACCTGGCCTACGACCTGTCCAAACATGCTGGCCATGCTCAACACGGCGCTCAACTTCTTC 250
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Qy 284 LeuTyrCysPheIleSerLysArgPheArg 293
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
Qy 251 CTGTACTGCTTCGTACGACGCGGTTCCGC 280
Db ||| : : : : : ||| : : : : : ||| : : : : : ||| : : : : : |||
```

Search completed: February 12, 2005, 15:53:21  
Job time : 3095 secs

**THIS PAGE BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:10:52 ; Search time 164 Seconds  
(without alignments)  
877.286 Million cell updates/sec

Title: US-10-689-832-20

Perfect score: 1967

Sequence: 1 MENTHAHLAANSLSWSPG.....NDSKSYQFEDAIGACVIL 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

1: Geneseqp1980s:\*

2: Geneseqp1990s:\*

3: Geneseqp2000s:\*

4: Geneseqp2001s:\*

5: Geneseqp2002s:\*

6: Geneseqp2003as:\*

7: Geneseqp2003bs:\*

8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1967	100.0	372	4	Aau10067 Chemokine
2	1967	100.0	372	5	Aae18645 Human G-p
3	1967	100.0	372	6	Abre2521 Human G-p
4	1967	100.0	372	7	Add18023 Human G-p
5	1967	100.0	372	7	Adj87767 G-coupled
6	1967	100.0	372	8	Adi79323 NOV10a pr
7	1962	99.7	372	8	Ado56002 Human NOV
8	1939.5	98.6	369	6	Abre2522 Human G-p
9	1864	94.8	591	7	Adf70485 Orphan re
10	1857	94.4	353	5	Aau76416 Novel G-p
11	1857	94.4	353	5	Abj04075 Human G p
12	1857	94.4	353	5	Abb79517 Human che
13	1857	94.4	353	6	Abp98724 Human axo
14	1857	94.4	353	6	Abra4437 Human G p
15	1857	94.4	353	6	Abb82499 Human TGR
16	1857	94.4	353	8	Adri16434 Human KOR
17	1853	94.2	353	4	Aau10068 Chemokine
18	1853	94.2	353	7	Adj87769 G-coupled
19	1853	94.2	353	8	Adi79325 NOV11 pro
20	1853	94.2	353	8	Ado56004 Human NOV
21	1853	94.2	353	8	Ado28954 Human nov
22	1829	93.0	353	5	Aae29236 Human 7TM
23	1829	93.0	353	6	Abg71163 Novel hum
24	1829	93.0	353	6	Abu09571 Human pro
25	1741	88.5	345	6	Abb82508 Mouse TGR

26	1741	88.5	345	8	ADO28956	Ado28956 Mouse nov
27	1628	82.8	318	7	ADJ87772	Adj87772 G-coupled
28	1628	82.8	318	7	ADI79348	Adi79348 NOV10b pr
29	1628	82.8	318	8	ADO56027	Ado56027 Human NOV
30	1623	82.5	333	5	AAE17081	Aae17081 Human G-p
31	1623	82.5	333	7	ABW00814	Abw00814 Human GPC
32	1623	82.5	343	5	ABB79519	Abb79519 Human che
33	1623	82.5	356	5	ABB79518	Abb79518 Human che
34	1623	82.5	385	5	ABP95606	Abp95606 Human GPC
35	1623	82.5	388	8	ADQ89946	Adq89946 Antagonis
36	1611	81.9	333	5	AAE17229	Aae17229 Human thy
37	1471	74.8	287	4	AAU25559	Aau25559 Human G p
38	1421	72.2	321	7	ADC12696	Adc12696 Human GPC
39	1317	67.0	265	6	ABP81706	Abp81706 Human G p
40	1317	67.0	265	6	ABR62524	Abre62524 Human G-p
41	912.5	46.4	313	4	AAU25556	Aau25556 Human G p
42	742	37.7	340	8	ADO28936	Ado28936 Mouse nov
43	742	37.7	348	6	ABR82431	Abre82431 Murine is
44	737	37.5	374	5	ABJ04070	Abj04070 Human G p
45	737	37.5	374	5	AAM49155	Aam49155 Human G p

ALIGNMENTS

RESULT 1

AAU10067

ID AAU10067 standard; protein; 372 AA.

AC AAU10067;

DT 14-FEB-2002 (first entry)

DE Chemokine receptor family related protein, NOV10.

XX NOV; cytostatic; psoriasis; nootropic; neuroprotectant;  
XX cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;  
XX haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;  
XX differentiation; proliferation; haematopoiesis; wound healing;  
XX angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;  
XX haemophilia; allergy; Penderg syndrome; skeletal dysplasia;  
XX ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;  
XX chemokine receptor; chromosome 1.  
XX Homo sapiens.

OS Homo sapiens.

PH Key Location/Qualifiers

FT Peptide 1..47

FT Protein /label= Signal\_peptide

XX 48..372

XX /note= "Mature chemokine receptor related protein, NOV10"

XX WO200170978-A2.

XX 27-SEP-2001.

XX 20-MAR-2001; 2001WO-US009093.

XX 20-MAR-2000; 2000US-0190768P.

XX 20-MAR-2000; 2000US-0190835P.

XX 22-MAR-2000; 2000US-0190972P.

XX 22-MAR-2000; 2000US-0191199P.

XX 24-MAR-2000; 2000US-0191947P.

XX 28-MAR-2000; 2000US-0192657P.

XX 28-MAR-2000; 2000US-0192664P.

XX 28-MAR-2000; 2000US-0192665P.

XX 28-MAR-2000; 2000US-0192984P.

XX 29-MAR-2000; 2000US-0192836P.

XX 31-MAR-2000; 2000US-0193843P.

XX (CURA-) CURAGEN CORP.

XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;

XX PI

PI Vernet CAM;  
XX WPI: 2001-639127/73.  
DR N-PSDB; AAS15730.  
XX  
PT Polypeptides and nucleic acids related to chloride channel, insulin-like  
PT growth factor family of proteins, useful for diagnosing and treating  
PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.  
XX  
XX Claim 1; Page 42; 151pp; English.  
XX  
CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX  
CC polypeptides are useful for treating pathology associated with NOVX  
CC polypeptide, determining the presence of or predisposition to a disease  
CC associated with altered levels of NOVX, identifying agents binding to  
CC NOVX and treatment of disorders associated with altered expression of  
CC members of chloride channel-associated proteins e.g. cystic fibrosis and  
CC congenital myotonia. NOVX proteins are useful in treatment of disorders  
CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,  
CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells  
CC in one or more organs (e.g. haemophilia, anaemia), Pendred syndrome,  
CC skeletal dysplasias, disorders characterised by altered cell shape,  
CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial  
CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the  
CC treatment of disorders of vascular smooth muscle cell differentiation,  
CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are  
CC useful to screen for molecules which inhibit or enhance NOVX activity or  
CC function and are useful as targets for the identifying small molecules,  
CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell  
CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV  
CC sequences are also useful for: identifying a cell or tissue type in a  
CC biological sample; amplifying DNA sequences from very small biological  
CC samples e.g. hair or skin or body fluids and as primers and probes to  
CC identify and/or clone NOVX homologues. NOVX proteins are useful  
CC immunogens to generate antibodies to monitor protein levels and modulate  
CC NOVX activity. Cells comprising the nucleic acids are useful for  
CC producing transgenic animals, for studying the function and/or activity  
CC of NOVX protein and identifying and/or evaluating modulators of NOVX  
CC protein activity. This sequence is the NOV10 amino acid sequence (gene  
CC located on chromosome 1) related to the chemokine receptor family of  
CC proteins, one of 12 NOV polypeptides described in the method of the  
CC invention  
XX  
SQ Sequence 372 AA;  
  
Query Match 100.0%; Score 1967; DB 4; Length 372;  
Best Local Similarity 100.0%; Pred. NO. 1.2e-209;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MEHTAHLAANSSLSWSPGSGACGLGFVPVYVYSLLCGLPANILTVIILSOLVARRQK 60  
DB 1 MEHTAHLAANSSLSWSPGSGACGLGFVPVYVYSLLCGLPANILTVIILSOLVARRQK 60  
  
QY 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVPDKIEVLEPSSHTSIWTV 120  
DB 61 SSYNYLLAALAAADILVLFVFDLLEDFILNMQMPQVPDKIEVLEPSSHTSIWTV 120  
  
QY 121 PLTIDRYIAVCHPLKHTVSVSPARTKVIIVYITCFLTSIPYWNPNWTEDYISTSVH 180  
DB 121 PLTIDRYIAVCHPLKHTVSVSPARTKVIIVYITCFLTSIPYWNPNWTEDYISTSVH 180  
  
QY 181 HVLINWCHFTVVLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTALFTITSFATL 240  
DB 181 HVLINWCHFTVVLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTALFTITSFATL 240  
  
QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLANTAINFLYCFISKRFRTAAATL 300  
DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLANTAINFLYCFISKRFRTAAATL 300  
  
QY 301 KAFFKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQDKNGKPIKSRNDSKSSVQ 360  
DB 301 KAFFKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQDKNGKPIKSRNDSKSSVQ 360

QY 361 FEDAIGACVIL 372  
DB 361 FEDAIGACVIL 372  
  
RESULT 2  
AAE18645  
ID AAE18645 standard; protein; 372 AA.  
XX  
AC AAE18645;  
XX  
DT 17-MAY-2002 (first entry)  
DE Human G-protein coupled receptor (GCRC-6).  
XX  
KW Human; G-protein coupled receptor; GCRC-6; cell proliferative disorder;  
KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
KW neurotic; neuroprotective; cardiant; immunosuppressive; anorectic;  
KW virucide; receptor.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Peptide 1..47  
FT Domain 32..52 /label= Signal\_peptide  
FT /note= "Transmembrane domain"  
FT Protein 48..372 /label= Human\_mature\_GCRC-6  
FT Domain 74..100 /note= "Transmembrane domain"  
FT Domain 184..208 /note= "Transmembrane domain"  
XX  
PN WO200210387-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 25-JUL-2001; 2001WO-US023433.  
XX  
PR 27-JUL-2000; 2000US-0221478P.  
PR 03-AUG-2000; 2000US-0223268P.  
PR 21-AUG-2000; 2000US-0227054P.  
PR 08-SEP-2000; 2000US-0231121P.  
PR 13-SEP-2000; 2000US-0232243P.  
PR 15-SEP-2000; 2000US-0232691P.  
PR 22-SEP-2000; 2000US-0235146P.  
XX  
PA (INCY-) INCYTE GENOMICS INC.  
XX  
PI Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
PI Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Wallia NK, Hafalia AJA;  
PI Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Graul RC;  
PI Warren BA, Lee BA, Ding L;  
XX  
DR WPI: 2002-188744/24.  
DR N-PSDB; AAD29672.  
XX  
PT New human G-protein coupled receptor polypeptide for diagnosis,  
PT prevention and treatment of cell proliferative, neurological,  
PT cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic  
PT disorders.  
XX  
PS Claim 1; Page 120; 150pp; English.  
XX  
CC The invention relates to novel human G-protein coupled receptors (GCRC)  
CC and their encoding polynucleotides. GCRC is useful as an immunogen for

CC preparing monoclonal and polyclonal antibodies. GCRC is useful for  
CC diagnosing, treating and preventing a cell proliferative disorder (e.g.,  
CC hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,  
CC Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's  
CC disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,  
CC myocardial infarction), gastrointestinal disorder (e.g., gastritis,  
CC cirrhosis, Crohn's disease), an autoimmune/ inflammatory disorder (e.g.,  
CC acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,  
CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,  
CC osteoporosis), and viral infections. GCRC is useful in a number of drug  
CC screening techniques, and to analyse the proteome of a tissue or cell  
CC type. GCRC is useful for creating knockin humanised animals or  
CC transgenic animals to model human diseases, in somatic or germline gene  
CC therapy, to generate a transcript image of a tissue or cell type, for  
CC detecting differences in the chromosomal location due to translocation,  
CC inversion, etc., among normal, carrier or affected individuals, and as  
CC hybridization probes for mapping naturally occurring genomic sequences.  
CC GCRC is useful in Southern or Northern analysis, dot blot or other  
CC membrane-based technologies, in PCR technologies, in dipstick, pin,  
CC multifunction enzyme linked immunosorbant (ELISA)-like assays, and in  
CC microarrays utilising fluids or tissues from patients to detect altered  
CC GCRC expression. The present sequence is human GCRC-6  
XX  
SQ Sequence 372 AA;

Query Match 100.0%; Score 1967; DB 5; Length 372;  
Best Local Similarity 100.0%; Pred. No. 1.2e-209; 0; Gaps 0;  
Matches 372; Conservative 0; Mismatches 0; Indels 0;

QY 1 MEHTAHLAANSLSWSPGACGLGFVPVYVYLLCLGLPANILTVILSOLVARRQK 60  
DB 1 MEHTAHLAANSLSWSPGACGLGFVPVYVYLLCLGLPANILTVILSOLVARRQK 60  
QY 61 SSNYLLALAAADILVLFVYVDFLEDFILNMQMPQDPDKLIEVLEPSSIHSTWTV 120  
DB 61 SSNYLLALAAADILVLFVYVDFLEDFILNMQMPQDPDKLIEVLEPSSIHSTWTV 120  
QY 121 PLTIDRYIAVCHPLKYHTVSPARTKRVISVYITCFLTSIPYWNPNWTEYISTSVH 180  
DB 121 PLTIDRYIAVCHPLKYHTVSPARTKRVISVYITCFLTSIPYWNPNWTEYISTSVH 180  
QY 181 HVLIIWHCFVYLVPCSIPIILNSIIVYKLRRKSNFLRGYSTGKTTAILFTITSFATL 240  
DB 181 HVLIIWHCFVYLVPCSIPIILNSIIVYKLRRKSNFLRGYSTGKTTAILFTITSFATL 240  
QY 241 WAPRIIMILYHLYGAPTONRWLHIMSDIANMLALTAINPFLYCFISKRPFTMAATL 300  
DB 241 WAPRIIMILYHLYGAPTONRWLHIMSDIANMLALTAINPFLYCFISKRPFTMAATL 300  
QY 301 KAFFKCKQKQPVQYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
DB 301 KAFFKCKQKQPVQYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
QY 361 FEDAIGACVUIL 372  
DB 361 FEDAIGACVUIL 372

RESULT 3

ABR62521  
ID ABR62521 standard; protein; 372 AA.  
XX  
AC ABR62521;  
XX  
DT 06-NOV-2003 (first entry)  
XX  
DE Human G-protein coupled receptor HGPBRMY34.  
XX  
KW HGPBRMY34; G-protein coupled receptor; GPCR-P14; GPCR-145;  
KW human; neuroprotective; nootropic; tranquilizer; antimigraine;  
KW neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;  
KW cytotatic; cardiant; hypotensive; antiangiinal; analgesic; anorectic;  
KW anti-HIV; antiasthmatic; osteopathic; uropathic; antiulcer; anti-allergic;  
KW

KW gene therapy.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Domain  
FT 16..36 Location/Qualifiers  
FT /label= TM1  
FT 65..87 /note= "transmembrane domain 1"  
FT /label= TM2  
FT /note= "transmembrane domain 2"  
FT 109..131 /label= TM3  
FT /note= "transmembrane domain 3"  
FT 148..166 /label= TM4  
FT /note= "transmembrane domain 4"  
FT 182..208 /label= TM5  
FT /note= "transmembrane domain 5"  
FT 227..249 /label= TM6  
FT /note= "transmembrane domain 6"  
FT 289..288 /label= TM7  
FT /note= "transmembrane domain 7"  
XX  
PN WO2003050256-A2.  
XX  
PD 19-JUN-2003.  
XX  
PF 06-DEC-2002; 2002WO-US039290.  
XX  
PR 06-DEC-2001; 2001US-0338371P.  
XX  
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.  
XX  
PI Feder JN, Gopal S, Mintier GA, Ramanathan CS;  
XX WPI: 2003-577295/54.  
DR N-PSDB; AC05275.  
XX  
PT New nucleic acid molecule encoding a human G-protein coupled receptor,  
PT HGPBRMY34, useful for diagnosing, preventing or treating diseases  
PT involving the receptor, for example Parkinson's disease, dementia,  
PT asthma, hypertension or cancer.  
XX  
XX Claim 5; Fig 1A-B; 112pp; English.

XX  
CC The present sequence is the protein sequence of human HGPBRMY34, a newly  
CC identified G-protein coupled receptor (GPCR) belonging to the group of  
CC 'Class A' GPCRs and showing homology to the Pfam model 7 transmembrane  
CC receptor, rhodopsin family. HGPBRMY34, also referred to as GPCR-P14  
CC and/or GPCR-145, is highly expressed in brain (especially in the  
CC amygdala, caudate nucleus, corpus callosum, hippocampus, thalamus,  
CC substantia nigra), spinal cord and pituitary, indicating an association  
CC in neurological systems and conditions. It is also expressed in the bone  
CC marrow and testis. The invention provides HGPBRMY34 polynucleotides,  
CC polypeptides and antibodies, expression vectors, host cells and antisense  
CC molecules, methods for screening for modulators of HGPBRMY34 activity  
CC and/or function, and methods for diagnosing, treating, preventing and  
CC screening for disorders and diseases associated with abnormal HGPBRMY34  
CC activity, including: a disorder related to aberrant G-protein coupled  
CC signalling; a disorder related to aberrant cell cycle regulation;  
CC neurological disorders; anxiety; headache; migraine; schizophrenia; manic  
CC depression; delirium; dementia; severe mental retardation and  
CC dyskinesias, such as Huntington's disease or Gilles de la Tourette's  
CC syndrome; Parkinson's disease; brain disorders; spinal cord disorders;  
CC affective disorders; neoplastic disorders; cardiovascular disorders;  
CC acute heart failure; hypotension; hypertension; angina pectoris;  
CC myocardial infarction; an immunological disorder; immune-related  
CC disorders; endocrinal diseases; growth disorders; neuropathic pain;  
CC obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis

CC ; psychosis; metabolic disorders; pituitary disorders; urinary retention;  
 CC ulcers; allergies; or benign prostatic hypertrophy (all claimed)  
 XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 1967; DB 6; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-209;  
 Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60  
 DB 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60  
 QY 61 SSNNYLLAALADILVLFVDFVDFLEDFILNMQMPQVDPKIIIEVLFSSHTSIWTV 120  
 DB 61 SSNNYLLAALADILVLFVDFVDFLEDFILNMQMPQVDPKIIIEVLFSSHTSIWTV 120  
 QY 121 PLTIDRYIAVCHPLKYHTVSPARTKVIIVSVYITCFLTSIPYYWPNWTEDYISTSVH 180  
 DB 121 PLTIDRYIAVCHPLKYHTVSPARTKVIIVSVYITCFLTSIPYYWPNWTEDYISTSVH 180  
 QY 181 HVLIIHCTVYLVPCSIFFILNSIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
 DB 181 HVLIIHCTVYLVPCSIFFILNSIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
 QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINFLYCFISKRPRTWAAATL 300  
 DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINFLYCFISKRPRTWAAATL 300  
 QY 301 KAFFKCKQKQVQFYTNHNFSTSSPISPANSHCHIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
 DB 301 KAFFKCKQKQVQFYTNHNFSTSSPISPANSHCHIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
 QY 361 FEDAIACVIL 372  
 DB 361 FEDAIACVIL 372

## RESULT 4

ADD18023  
 ID ADD18023 standard; protein; 372 AA.

XX  
 AC ADD18023;

XX  
 DT 15-JAN-2004 (first entry)

XX  
 DE Human G-protein coupled receptor (GPCR) sequence Seq ID19.

XX  
 KW G protein coupled receptor; GPCR; signal transduction pathway; G protein;  
 KW Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;  
 KW colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;  
 KW sleeplessness; hypertension; anxiety; stress; renal failure;  
 KW cardiovascular disorder; neural disorder; oncology disorder;  
 KW immune disorder; neuroprotective; gene therapy.

XX  
 OS Homo sapiens.

XX  
 FH Key Location/Qualifiers

FT Misc-difference 347.351  
 FT /note= "Encoded by AAGTC"

XX  
 FN WO2003016478-A2.

XX  
 PD 27-FEB-2003.

XX  
 PF 15-AUG-2002; 2002WO-US026017.

XX  
 PR 20-AUG-2001; 2001US-0313658P.

XX  
 PR 12-SEP-2001; 2001US-0318675P.

XX  
 PR 30-OCT-2001; 2001US-0340703P.

XX  
 PR 26-NOV-2001; 2001US-0333417P.

XX  
 PR 06-DEC-2001; 2001US-0338367P.

XX  
 PR 06-FEB-2002; 2002US-0355596P.

XX  
 PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX  
 PI Feder JN, Ramanathan CS, Gopal S, Mintier GA;

XX  
 DR WPI; 2003-278558/27.

XX  
 DR N-PSDB; ADD18120.

XX  
 PT New nucleic acid, useful for manufacturing a medicament for preventing,  
 PT treating or ameliorating a medical condition e.g., neural disorder.

XX  
 PS Example 1; SEQ ID NO 19; 251pp; English.

XX  
 CC This invention relates to novel G protein coupled receptors (GPCRs) and  
 CC their encoding nucleotide sequences. Many medically significant  
 CC biological processes are mediated by proteins participating in signal  
 CC transduction pathways involving G proteins. GPCRs are one of the largest  
 CC receptor superfamilies known. These receptors are biologically important  
 CC and malfunction of these receptors results in diseases such as  
 CC Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal  
 CC pigmentosa and asthma. They are also involved in depression,  
 CC schizophrenia, sleeplessness, hypertension, anxiety, stress, renal  
 CC failure and other cardiovascular, neural, oncology and immune disorders.  
 CC A modulator of the GPCRs of the invention may have neuroprotective  
 CC activity whilst the sequences of the invention may be useful for gene  
 CC therapy. The invention may also be useful for manufacturing a medicament  
 CC for preventing, treating or ameliorating a medical condition. The present  
 CC sequence is the amino acid sequence of a human GPCR of the invention.

XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 1967; DB 7; Length 372;

Best Local Similarity 100.0%; Pred. No. 1.2e-209; Indels 0; Gaps 0;

Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60

DB 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60

QY 61 SSNNYLLAALADILVLFVDFVDFLEDFILNMQMPQVDPKIIIEVLFSSHTSIWTV 120

DB 61 SSNNYLLAALADILVLFVDFVDFLEDFILNMQMPQVDPKIIIEVLFSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKYHTVSPARTKVIIVSVYITCFLTSIPYYWPNWTEDYISTSVH 180

DB 121 PLTIDRYIAVCHPLKYHTVSPARTKVIIVSVYITCFLTSIPYYWPNWTEDYISTSVH 180

QY 181 HVLIIHCTVYLVPCSIFFILNSIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240

DB 181 HVLIIHCTVYLVPCSIFFILNSIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINFLYCFISKRPRTWAAATL 300

DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINFLYCFISKRPRTWAAATL 300

QY 301 KAFFKCKQKQVQFYTNHNFSTSSPISPANSHCHIKMLVYQYDKNGKPIKSRNDSKSSYQ 360

DB 301 KAFFKCKQKQVQFYTNHNFSTSSPISPANSHCHIKMLVYQYDKNGKPIKSRNDSKSSYQ 360

QY 361 FEDAIACVIL 372

DB 361 FEDAIACVIL 372

## RESULT 5

ADJ87767

ID ADJ87767 standard; protein; 372 AA.

XX  
 AC ADJ87767;

XX  
 DT 06-MAY-2004 (first entry)

XX  
 DE G-coupled protein receptor-related protein #56.

XX novel protein; G-coupled protein receptor-related protein;  
 KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;  
 KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;  
 KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;  
 KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;  
 KW viral infections; bacterial infection; parasitic infection;  
 KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;  
 KW Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2002102321-A2.  
 XX  
 PD 27-DEC-2002.  
 XX  
 XX 18-JUN-2002; 2002WO-US019522.  
 XX  
 PR 18-JUN-2001; 2001US-0298994P.  
 PR 18-JUN-2001; 2001US-0299134P.  
 PR 04-OCT-2001; 2001US-00972446.  
 PR 06-JUN-2002; 2002US-00299134.  
 PR 07-JUN-2002; 2002US-00298994.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PG;  
 PI Pena CE, Spaderna SK, Zhong M;  
 XX  
 DR WPI: 2003-167441/16.  
 DR N-PSDB; ADJ87766.  
 XX  
 PT New MOLX polypeptides and polynucleotides, useful in gene therapy,  
 PT particularly for treating or preventing e.g. cardiomyopathy,  
 PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate  
 PT cancer, stroke or pain.  
 XX  
 XX Claim 1; SEQ ID NO 202; 378pp; English.  
 XX  
 CC The invention comprises the amino acid and coding sequences of novel G-  
 CC coupled protein receptor-related (MOL) proteins. The DNA and protein  
 CC sequences of the invention are useful for treating or preventing a MOL-  
 CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders  
 CC associated with cell signal processing and metabolic pathway modulation,  
 CC or diabetes. The DNA and protein sequences are also useful for the  
 CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,  
 CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,  
 CC systemic lupus erythematosus, viral infections, bacterial infections,  
 CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau  
 CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or  
 CC cerebral palsy. The present amino acid sequence represents a MOL protein  
 CC of the invention.  
 XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 1967; DB 7; Length 372;  
 Best Local Similarity 100.0%; Pred. No. 1.2e-209;  
 Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLWSFGSACGLGFPVVPVYSLILCLGLPANILTVILSOLVARQK 60  
 DB |||||||  
 DB 1 MEHTHAHLAANSLWSFGSACGLGFPVVPVYSLILCLGLPANILTVILSOLVARQK 60  
 QY 61 SSNYLLAALAAADILVLFVTFVDFLEDPILNMQMPQVDKILVLEFSSIHSTWITV 120  
 DB |||||||  
 DB 61 SSNYLLAALAAADILVLFVTFVDFLEDPILNMQMPQVDKILVLEFSSIHSTWITV 120  
 QY 121 PLTIDRYIACHPLKHYTVSPARTKRVISVYITCFLTSIPYWPNIWTEYISTSVH 180  
 DB |||||||  
 DB 121 PLTIDRYIACHPLKHYTVSPARTKRVISVYITCFLTSIPYWPNIWTEYISTSVH 180  
 QY 181 HVLIIWHCTFVYLVPSCIFFILNSIIYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240  
 DB |||||||

DB 181 HVLIIWHCTFVYLVPSCIFFILNSIIYKLRKSNFRLRGYSTGKTTAILFTTISIPATL 240  
 QY 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRFRMTAAATL 300  
 DB |||||||  
 DB 241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRFRMTAAATL 300  
 QY 301 KAFFKCKQKQVQPYTHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
 DB |||||||  
 DB 301 KAFFKCKQKQVQPYTHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
 QY 361 FEDAIGACVIL 372  
 DB |||||||  
 DB 361 FEDAIGACVIL 372  
 RESULT 6  
 ADI79323  
 ID ADI79323 standard; protein; 372 AA.  
 XX  
 AC ADI79323;  
 XX  
 DT 22-APR-2004 (first entry)  
 XX  
 DE NOV10a protein sequence, SEQ ID 20.  
 XX  
 KW Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;  
 KW Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;  
 KW Ophthalmological; Antipsoriatic; Neuroprotective; Nootropic;  
 KW Antiarteriosclerotic; Hypotensive; Cardiant; Cerebroprotective;  
 KW Gene Therapy; NOVX; human; cancer; myelogenous leukaemia;  
 KW congenital neonatal autoimmune thrombocytopaenia; immunological disorder;  
 KW allergy; infection; asthma; lung disease; reproductive disorder;  
 KW haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;  
 KW diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;  
 KW hypertension; stroke; heart failure; chromosome 1; NOV10a;  
 KW chemokine receptor.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2004009635-A2.  
 XX  
 PD 29-JAN-2004.  
 XX  
 PF 04-OCT-2001; 2001WO-US031292.  
 XX  
 PR 20-MAR-2001; 2001US-00813432.  
 XX  
 PA (CURA-) CURAGEN CORP.  
 XX  
 XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;  
 PI Vernet CAM;  
 XX  
 DR WPI: 2004-123380/12.  
 DR N-PSDB; ADI79322.  
 XX  
 XX Claim 1; Page 41; 158pp; English.  
 XX  
 CC The present invention relates to novel NOVX proteins and their coding  
 CC sequences (ADI79304-ADI79327). The sequences are useful for the  
 CC manufacture of a medicament for treating a syndrome associated with a  
 CC human disease associated with the protein, or for diagnosing and treating  
 CC disorders associated with the NOVX protein, such as cancer, myelogenous  
 CC leukaemia, congenital neonatal autoimmune thrombocytopaenia,  
 CC immunological disorders, allergy and infection, asthma, lung diseases,  
 CC reproductive disorders, male and female reproductive diseases,  
 CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,  
 CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,  
 CC hypertension, stroke and heart failure. NOV10a represents a new member of  
 CC the chemokine receptor family and the gene is located on chromosome 1.  
 XX  
 SQ Sequence 372 AA;

Query Match 100.0%; Score 1967; DB 8; Length 372;

Best Local Similarity 100.0%; Pred. No. 1.2e-209;		CC invention.	
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		XX	SQ
		Sequence 372 AA;	
QY	1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSSLLCLGLPANILTVIILSOLVARQK 60	Query Match 99.7%; Score 1962; DB 8; Length 372;	
Db	1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSSLLCLGLPANILTVIILSOLVARQK 60	Best Local Similarity 99.7%; Pred. No. 4.2e-209;	
		Matches 371; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	61 SSYNVLLAALAAADILVLPFIVDFVDFLEDFILNMQMPQVPDKIIIEVLFSSHTSIWITV 120	QY	1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSSLLCLGLPANILTVIILSOLVARQK 60
Db	61 SSYNVLLAALAAADILVLPFIVDFVDFLEDFILNMQMPQVPDKIIIEVLFSSHTSIWITV 120	Db	1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSSLLCLGLPANILTVIILSOLVARQK 60
QY	121 PLTIDRYIAVCHPLKYHTVSYPARTRKVIIVSYITCFLTSIPYVWPNWTEDYISTSVH 180	QY	61 SSYNVLLAALAAADILVLPFIVDFVDFLEDFILNMQMPQVPDKIIIEVLFSSHTSIWITV 120
Db	121 PLTIDRYIAVCHPLKYHTVSYPARTRKVIIVSYITCFLTSIPYVWPNWTEDYISTSVH 180	Db	61 SSYNVLLAALAAADILVLPFIVDFVDFLEDFILNMQMPQVPDKIIIEVLFSSHTSIWITV 120
QY	181 HVLWIHCFTVYLPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTALFTITSIFATL 240	QY	121 PLTIDRYIAVCHPLKYHTVSYPARTRKVIIVSYITCFLTSIPYVWPNWTEDYISTSVH 180
Db	181 HVLWIHCFTVYLPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTALFTITSIFATL 240	Db	121 PLTIDRYIAVCHPLKYHTVSYPARTRKVIIVSYITCFLTSIPYVWPNWTEDYISTSVH 180
QY	241 WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALNTAINFFLYCFISKRPRTWAAATL 300	QY	181 HVLWIHCFTVYLPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTALFTITSIFATL 240
Db	241 WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALNTAINFFLYCFISKRPRTWAAATL 300	Db	181 HVLWIHCFTVYLPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTALFTITSIFATL 240
QY	301 KAFFKQKQPVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360	QY	241 WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALNTAINFFLYCFISKRPRTWAAATL 300
Db	301 KAFFKQKQPVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360	Db	241 WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALNTAINFFLYCFISKRPRTWAAATL 300
QY	361 FEDAIGACVIL 372	QY	301 KAFFKQKQPVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360
Db	361 FEDAIGACVIL 372	Db	301 KAFFKQKQPVQFYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360
RESULT 7		QY	361 FEDAIGACVIL 372
ADOS6002		Db	361 FEDAIGACVIL 372
ID ADOS6002 standard; protein; 372 AA.		RESULT 8	
AC	ADOS6002;	ABR62522	
XX		ID ABR62522 standard; protein; 369 AA.	
DT	15-JUL-2004 (first entry)	XX	AC ABR62522;
DE	Human NOV10 polypeptide.	XX	DT 06-NOV-2003 (first entry)
KW	human; cancer; obesity; diabetic nephropathy; acute pancreatitis; stroke;	XX	DE Human G-protein coupled receptor HGPBMY34 variant.
KW	multiple sclerosis.	KW	HGPBMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;
XX		KW	human; neuroprotective; nootropic; tranquilizer; antimigraine;
OS	Homo sapiens.	KW	neuroleptic; antidmanic; antidepressant; anticonvulsant; antiparkinsonian;
PN	US2004058862-A1.	KW	cytostatic; cardiant; hypotensive; antitanginal; analgesic; anorectic;
PD	25-MAR-2004.	KW	anti-HIV; antiasthmatic; osteopathic; uropathic; antiulcer; antiallergic;
XX		XX	gene therapy.
PF	18-SEP-2002; 2002US-00246593.	OS	Homo sapiens.
XX		XX	
PR	18-SEP-2002; 2002US-00246583.	XX	Key
XX		XX	Location/Qualifiers
PA	(MAJU/) MAJUMDER K.	XX	16. .37
XX		XX	/label= TM1
PI	Majumder K;	XX	/note= "transmembrane domain 1"
XX		XX	65. .87
DR	WPI; 2004-268835/25.	XX	/label= TM2
DR	N-PSDB; ADO56001.	XX	/note= "transmembrane domain 2"
XX		XX	109. .131
XX		XX	/label= TM3
XX		XX	/note= "transmembrane domain 3"
XX		XX	148. .166
XX		XX	/label= TM4
XX		XX	/note= "transmembrane domain 4"
XX		XX	182. .208
XX		XX	/label= TM5
XX		XX	/note= "transmembrane domain 5"
XX		XX	227. .249
XX		XX	
CC	The invention relates to novel isolated NOVX nucleic acids and encoded	XX	
CC	polypeptides. The nucleic acids, polypeptides and antibodies raised	XX	
CC	against the polypeptides are useful for preventing or treating diseases	XX	
CC	associated with aberrant NOVX expression or activity e.g., cancer,	XX	
CC	obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple	XX	
CC	sclerosis. The present sequence represents a NOVX polypeptide of the	XX	



FT /label= TM6  
 FT /note= "transmembrane domain 6"  
 FT 269. .288  
 FT /label= TM7  
 FT /note= "transmembrane domain 7"  
 FT 19-JUN-2003.  
 PN WO2003050256-A2.  
 XX 19-JUN-2003.  
 XX 06-DEC-2002; 2002WO-US039290.  
 XX 06-DEC-2001; 2001US-0338371P.  
 PR (BRIM ) BRISTOL-MYERS SQUIBB CO.  
 PA Feder JN, Gopal S, Mintier GA, Ramanathan CS;  
 PI WPI; 2003-577295/54.  
 DR N-PSDB; ACF05276.  
 XX  
 XX New nucleic acid molecule encoding a human G-protein coupled receptor,  
 PT HGPBRMY34, useful for diagnosing, preventing or treating diseases  
 PT involving the receptor, for example Parkinson's disease, dementia,  
 PT asthma, hypertension or cancer.  
 XX  
 PS Claim 5; Fig 2A-B; 112pp; English.  
 XX  
 CC The present sequence is the protein sequence of a variant of human  
 CC HGPBRMY34 (see AB862521) containing a deletion of 3 amino acids but  
 CC expected to share at least some of the expression patterns and function  
 CC of HGPBRMY34. HGPBRMY34 is a newly identified G-protein coupled receptor  
 CC (GPCR) belonging to the group of 'Class A' GPCRs and showing homology to  
 CC the Pfam model 7 transmembrane receptor, rhodopsin family. HGPBRMY34 (or  
 CC GPCR-P14, GPCR-145) is highly expressed in brain (amygdala, caudate  
 CC nucleus, corpus callosum, hippocampus, thalamus, substantia nigra),  
 CC spinal cord and pituitary, indicating an association in neurological  
 CC systems and conditions. It is also expressed in the bone marrow and  
 CC testis. The invention provides HGPBRMY34 polynucleotides, polypeptides  
 CC and antibodies, expression vectors, host cells and antisense molecules,  
 CC methods for screening for modulators of HGPBRMY34 activity and/or  
 CC function, and methods for diagnosing, treating, preventing and screening  
 CC for disorders and diseases associated with abnormal HGPBRMY34 activity,  
 CC including: a disorder related to aberrant G-protein coupled signalling; a  
 CC disorder related to aberrant cell cycle regulation; neurological  
 CC disorders; anxiety; headache; migraine; schizophrenia; manic depression;  
 CC delirium; dementia; severe mental retardation and dyskinesias such as  
 CC Huntington's disease or Gilles de la Tourette's syndrome; Parkinson's  
 CC disease; brain disorders; spinal cord disorders; affective disorders;  
 CC neoplastic disorders; cardiovascular disorders; acute heart failure;  
 CC hypotension; hypertension; angina pectoris; myocardial infarction; an  
 CC immunological disorder; immune-related disorders; endocrinal diseases;  
 CC growth disorders; neuropathic pain; obesity; anorexia; HIV infections;  
 CC cancers; bulimia; asthma; osteoporosis; psychosis; metabolic disorders;  
 CC pituitary disorders; urinary retention; ulcers; allergies; or benign  
 CC prostatic hypertrophy (all claimed)  
 XX  
 XX Sequence 369 AA;  
 SQ  
 Query Match 98.6%; Score 1939.5; DB 6; Length 369;  
 Best Local Similarity 99.2%; Pred. No. 1.3e-206;  
 Matches 369; Conservative 0; Mismatches 0; Indels 3; Gaps 1;  
 QY 1 MEHTHAHLAANSLSWNSPGSACGLGFPVVPVYVYSLLLCLGLPANILTVILSGLVARRKK 60  
 DB 1 MEHTHAHLAANSLSWNSPGSACGLGFPVVPVYVYSLLLCLGLPANILTVILSGLVARRKK 60  
 QY 61 SSNNYLLAALADILVLFVFDLLEDFFILNMQMPQVPDKIIIEVLFSSHTSIWTV 120  
 DB 61 SSNNYLLAALADILVLFVFDLLEDFFILNMQMPQVPDKIIIEVLFSSHTSIWTV 120  
 QY 121 PLTIDRYIACHPLKHTVSVYPARTKVIUSVITCFLTSIPYVWPNWEDYISTSVH 180  
 DB 121 PLTIDRYIACHPLKHTVSVYPARTKVIUSVITCFLTSIPYVWPNWEDYISTSVH 180

Db 121 PLTIDRYIACHPLKHTVSVYPARTKVIUSVITCFLTSIPYVWPNWEDYISTSVH 180  
 QY 181 HVLIIWHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTTISIPATL 240  
 Db 181 HVLIIWHCFVTVLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTATLFTTISIPATL 240  
 QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINLPFLYCFISKRFRTMAAATL 300  
 Db 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLLNTAINLPFLYCFISKRFRTMAAATL 300  
 QY 301 KAFFKCKQKOPVOFYTNHNFSSITSSPWI SPANSHCIKMLVVOYDKNGKPKSRNDSKSSYQ 360  
 Db 301 KAFFKCKQKOPVOFYTNHNFSSITSSPWI SPANSHCIKMLVVOYDKNGKPKSRNDSKSSYQ 357  
 QY 361 FEDAIGACVIL 372  
 Db 358 FEDAIGACVIL 369  
 RESULT 9  
 ADF70485  
 ID ADF70485 standard; protein; 591 AA.  
 XX  
 AC ADF70485;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Orphan receptor ligand-related human protein SeqID108.  
 XX  
 KW ligand; orphan receptor protein; fusion protein; fluorescent protein;  
 KW cell expression; green fluorescent protein; GFP; GFP-1; wild-type GFP;  
 KW GFPuv; Enhanced GFP; EGFP; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003071272-A1.  
 XX  
 PD 28-AUG-2003.  
 XX  
 PF 21-FEB-2003; 2003WO-JP001901.  
 XX  
 PR 22-FEB-2002; 2002JP-00045728.  
 PR 23-JUL-2002; 2002JP-00213949.  
 PR 11-OCT-2002; 2002JP-00298237.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Hinuma S, Fujii R, Ogi K, Komateu H, Kawamata Y, Hosoya M;  
 XX  
 DR WPI; 2003-697654/66.  
 DR N-PSDB; ADF70587.  
 XX  
 PT Transformation of cells with a fusion protein of an orphan receptor  
 PT protein with a fluorescent protein useful for identification of ligands  
 PT to the orphan receptor.  
 XX  
 PS Disclosure; SEQ ID NO 108; 594pp; Japanese.  
 XX  
 CC This invention relates to a novel method of identifying ligands to an  
 CC orphan receptor protein which comprises transforming cells with DNA  
 CC encoding a fusion protein of the orphan receptor with a fluorescent  
 CC protein, so that the fusion protein is expressed in the cells (or cell  
 CC membranes isolated from them) and contacting the cells with the potential  
 CC ligand to be tested. A suitable fluorescent protein for incorporation in  
 CC the fusion protein is green fluorescent protein (GFP), for example GFP-1,  
 CC wild-type GFP, GFPuv or Enhanced GFP (EGFP). The method is useful for the  
 CC identification of ligands binding to an orphan receptor protein.  
 XX  
 SQ Sequence 591 AA;  
 Query Match 94.8%; Score 1864; DB 7; Length 591;  
 Best Local Similarity 96.7%; Pred. No. 6.5e-198;  
 Matches 353; Conservative 1; Mismatches 11; Indels 0; Gaps 0;

Qy	1	MEHTHA	LAANSSLSWSPGSACGLGFVPVVVYSLLLCLGLPANILTVIISLVARRQK	60
Dd	1	MEHTHA	LAANSSLSWSPGSACGLGFVPVVVYSLLLCLGLPANILTVIISLVARRQK	60
Qy	61	SSNYL	LALAAADILVLFFIVFVDPLEBDFILNMQPQDPDKIEVLFPSSIHTSIWTV	120
Dd	61	SSNYL	LALAAADILVLFFIVFVDPLEBDFILNMQPQDPDKIEVLFPSSIHTSIWTV	120
Qy	121	PLTIDRYIAVCHPLKYHTVSYPARTRKVIUSVYTCTELTSIPYYMWFNIWTEDIYSTSVH	180	
Dd	121	PLTIDRYIAVCHPLKYHTVSYPARTRKVIUSVYTCTELTSIPYYMWFNIWTEDIYSTSVH	180	
Qy	181	HVLWIWHCFVTYLVPCSIFFILNSTIIYVKLRKSNFRLGYSTGKTALLFTITTSIFATL	240	
Dd	181	HVLWIWHCFVTYLVPCSIFFILNSTIIYVKLRKSNFRLGYSTGKTALLFTITTSIFATL	240	
Qy	241	WAPRIIMILYHGAPIQNRMVLVHMDSIANMLALTAINFLFYCFISKRPRTMAAATL	300	
Dd	241	WAPRIIMILYHGAPIQNRMVLVHMDSIANMLALTAINFLFYCFISKRPRTMAAATL	300	
Qy	301	KAFFKCQKQPVQFYTNHNFSITSSPWISPANSCHIKMLVYQDKNKGPKIKSRNDSKSXY	360	
Dd	301	KAFFKCQKQPVQFYTNHNFSITSSPWISPANSCHIKMLVYQDKNKGPKIKSVSSSGEEL	360	
Qy	361	FEDAI	365	
Dd	361	FTGVG	365	:

RESULT 10	
AAU76416	
ID	AAU76416 standard; protein; 353 AA.
XX	
AC	AAU76416;
XX	
XX	
DT	08-MAY-2002 (first entry)
XX	
XX	
XX	Novel G-protein coupled receptor TGR8.
XX	
KW	G-protein coupled; receptor; cardiovascular; immunomodulator; TGR8;
KW	cytostatic; antiinflammatory; antiulcer; fetal brain;
KW	central nervous system disease; circulatory organ disorder; cancer;
KW	metabolic disease; immunological disease; gastrointestinal disease;
KW	gene therapy; transgenic animal; human.
XX	
OS	Homo sapiens.
XX	
XX	WO200194582-A1.
XX	
PD	13-DEC-2001.
XX	
XX	01-JUN-2001; 2001WO-JP004643.
XX	
PR	02-JUN-2000; 2000JP-00170446.
PR	23-JUN-2000; 2000JP-00194926.
XX	
XX	(TAKE ) TAKEDA CHEM IND LTD.
PA	
XX	
PI	Terao Y, Matsui H, Shintani Y;
XX	
DR	WPI: 2002-164317/21.
DR	N-PSDB; ABK15562, ABK15563.
XX	
XX	Human fetal brain-originated G protein-coupled receptor protein TGR8 and
PT	encoding DNA, for developing drugs to treat e.g. diseases of the central
PT	nervous system or circulatory organs, cancer, and metabolic diseases.
XX	
XX	Claim 1; Fig 2; 102pp; Japanese.
PS	
XX	
CC	The invention describes a human fetal brain-originated G protein-coupled
CC	receptor protein, or its salt. The protein and encoded DNA are useful for
CC	developing drugs to treat e.g. diseases of the central nervous system or

circulatory organs, cancer, metabolic diseases, immunological diseases and gastrointestinal diseases. The invention also describes creation of a probe for gene therapy and construction of a transgenic animal. This is the amino acid sequence of the novel G-protein coupled receptor TGR8, described in the method of the invention

Query Match	94.4%	Score 1857;	DB 5;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 1.9e-197;		
Matches 350; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	MEHTAHLAANSLSWMSPGSACGLGPPVPVYYSLLCLGLPANILTVIILSQLVARQK	60	
Db	1	MEHTAHLAANSLSWMSPGSACGLGPPVPVYYSLLCLGLPANILTVIILSQLVARQK	60	
Qy	61	SSNYLLALAAADILVLFVFDLLEDFILNMQPQVPDKIIEVLEFSSIHTSITV	120	
Db	61	SSNYLLALAAADILVLFVFDLLEDFILNMQPQVPDKIIEVLEFSSIHTSITV	120	
Qy	121	PLTIDRYIACVCHPLKHYHTVSPYARTRKIVSVYITCTFLTSPYVWPNWITVEDYISTSVH	180	
Db	121	PLTIDRYIACVCHPLKHYHTVSPYARTRKIVSVYITCTFLTSPYVWPNWITVEDYISTSVH	180	
Qy	181	HVLIIWHCFVYLVPCSIIFILANSIIYVKLRKSNFRLRGYSTCKTTAILFTITSIPATL	240	
Db	181	HVLIIWHCFVYLVPCSIIFILANSIIYVKLRKSNFRLRGYSTCKTTAILFTITSIPATL	240	
Qy	241	WAPRIIMILYHLYGAPIQNERLWVHMSDIANMLALNTAINFFLYCTISKFRFRTMAAATL	300	
Db	241	WAPRIIMILYHLYGAPIQNERLWVHMSDIANMLALNTAINFFLYCTISKFRFRTMAAATL	300	
Qy	301	KAPFKCQKQPVQFYTNHNFISITSPWITSPANSHCIKMLVQYDKNKGPIK	350	
Db	301	KAPFKCQKQPVQFYTNHNFISITSPWITSPANSHCIKMLVQYDKNKGPIK	350	

RESULT 11	
ABJ04075	
ID	ABJ04075 standard; protein; 353 AA.
XX	
AC	ABJ04075;
XX	
DT	11-OCT-2002 (first entry)
XX	
DE	Human G protein coupled receptor hRUP35.
XX	
KW	Human; G-protein coupled receptor; GPCR; hRUP28; hRUP29; hRUP30; hRUP31;
KW	hRUP32; hRUP33; hRUP34; hRUP35; hRUP36; hRUP37.
XX	
OS	Homo sapiens.
XX	
WO	WO200242461-A2.
XX	
PN	30-MAY-2002.
PD	
XX	
PF	26-NOV-2001; 2001WO-US044386.
XX	
PR	27-NOV-2000; 2000US-0253404P.
PR	12-DEC-2000; 2000US-0255366P.
PR	20-FEB-2001; 2001US-0370266P.
PR	20-FEB-2001; 2001US-0270286P.
PR	06-APR-2001; 2001US-0282032P.
PR	06-APR-2001; 2001US-0282356P.
PR	06-APR-2001; 2001US-0282358P.
PR	06-APR-2001; 2001US-0282365P.
PR	14-MAY-2001; 2001US-0290917P.
PR	31-JUL-2001; 2001US-0309208P.
XX	
PA	(AREN-) ARENA PHARM INC.
XX	
PI	Chen R, Chu ZL, Dang HT, Lowitz KP, Pride C;
XX	

DR WPI: 2002-566565/60.  
 XX N-PSDB; ABT04873.  
 PT Novel endogenous and non-endogenous versions of G protein-coupled  
 PT receptor useful for identification of candidate compounds as receptor  
 PT agonists or antagonists for use as therapeutic agents.  
 XX  
 PS Claim 29; Page 70-72; 84pp; English.  
 CC  
 CC The present invention provides the protein and coding sequences of  
 CC several human G-protein coupled receptors (GPCRs). These can be used in  
 CC the identification of candidate compounds as receptor agonists or inverse  
 CC agonists having applicability as therapeutic agents. The present sequence  
 CC is a GPCR protein of the invention  
 XX  
 SQ Sequence 353 AA;

Query Match 94.4%; Score 1857; DB 5; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-197;  
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLLCLGLPANILTVILSOLVARROK 60  
 DB 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLLCLGLPANILTVILSOLVARROK 60  
 QY 61 SSYNLLALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSIHSTWIV 120  
 DB 61 SSYNLLALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSIHSTWIV 120  
 QY 121 PLTIDRYIAVCHPLKHTVTSYPARTKRVISVYITCFLTSIPYVWPNWTEYISTSVH 180  
 DB 121 PLTIDRYIAVCHPLKHTVTSYPARTKRVISVYITCFLTSIPYVWPNWTEYISTSVH 180  
 QY 181 HVLWIHCHFTVYLVPSCIPFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240  
 DB 181 HVLWIHCHFTVYLVPSCIPFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240  
 QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFLYCFISKRFPTMAAATL 300  
 DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFLYCFISKRFPTMAAATL 300  
 QY 301 KAFFKCKQKQPVQYTNHNFSSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350  
 DB 301 KAFFKCKQKQPVQYTNHNFSSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350

## RESULT 12

ABB79517  
 ID ABB79517 standard; protein; 353 AA.

XX  
 AC ABB79517;  
 XX  
 DT 23-SEP-2002 (first entry)  
 XX  
 DE Human chemokine-like receptor.  
 XX  
 KW Chemokine-like receptor; G-protein coupled receptor; receptor; human;  
 KW HIV infection; cardiovascular disease; asthma;  
 KW chronic obstructive pulmonary disease; cardiac; antiasthmatic;  
 KW vasotropic; hypotensive; antiarrhythmic; thrombolytic; anticoagulant;  
 KW antiinflammatory; antiallergic; immunomodulator; gene therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 OS WO200248358-A2.  
 PN  
 PD 20-JUN-2002.  
 XX  
 PF 12-DEC-2001; 2001WO-EP014571.  
 XX  
 PR 14-DEC-2000; 2000US-0255150P.  
 PR 02-APR-2001; 2001US-0280110P.  
 PR 21-JUN-2001; 2001US-0299474P.

XX  
 PA (FARB ) BAYER AG.  
 XX  
 PI Smolyar A, Zhu Z, Encinas J, Watanabe S, Okigami H;  
 XX  
 DR WPI: 2002-547858/58.  
 DR N-PSDB; ABN84269, ABN84273.  
 XX  
 PT New isolated polynucleotide encoding a chemokine-like receptor  
 PT polypeptide for treating e.g. asthma, myocardial infarction, human  
 PT immunodeficiency virus infection, or chronic obstructive pulmonary  
 PT disease.  
 XX  
 XX Claim 25; Fig 2; 114pp; English.

XX  
 CC The present sequence is the protein sequence of a novel human chemokine-  
 CC like receptor of 353 amino acids. The chemokine-like receptor has 7  
 CC putative transmembrane domains, consistent with the structure of a G-  
 CC protein coupled receptor. Its closest human homologue is C-C chemokine  
 CC receptor 3. The novel receptor is expressed at low levels in most  
 CC tissues. It is expressed at a high level in phytohaemagglutinin-  
 CC stimulated CD8+ cells, but in none of the other immune cells tested. It  
 CC may act as a receptor of chemoattractant molecules on activated  
 CC lymphocytes and be involved in cell trafficking and homing to sites of  
 CC infection, inflammation or tissue injury. Regulation of activity of the  
 CC novel receptor can therefore be used to treat cardiovascular,  
 CC immunological and inflammatory diseases, including asthma and chronic  
 CC obstructive pulmonary disease (COPD). The receptor may also be a target  
 CC for viruses that reside in the nervous system. Regulating the binding of  
 CC ligands, e.g. chemoattractant molecules or virus particles, to the  
 CC receptor can therefore be used to modulate the immune response to inhibit  
 CC viral infections, including HIV infection. A claimed method of reducing  
 CC activity of an antibody, antisense oligonucleotide or ribozyme) to a  
 CC product (preferably RNA or a polypeptide) encoded by a polynucleotide  
 CC encoding the human chemokine-like receptor in vivo or in vitro. A claimed  
 CC method of treating a chemokine-like receptor dysfunction related disease  
 CC selected from HIV infection, a cardiovascular disorder, asthma or COPD  
 CC uses a reagent that modulates a function of the receptor  
 XX  
 SQ Sequence 353 AA;

Query Match 94.4%; Score 1857; DB 5; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-197;  
 Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLLCLGLPANILTVILSOLVARROK 60  
 DB 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLLCLGLPANILTVILSOLVARROK 60  
 QY 61 SSYNLLALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSIHSTWIV 120  
 DB 61 SSYNLLALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSIHSTWIV 120  
 QY 121 PLTIDRYIAVCHPLKHTVTSYPARTKRVISVYITCFLTSIPYVWPNWTEYISTSVH 180  
 DB 121 PLTIDRYIAVCHPLKHTVTSYPARTKRVISVYITCFLTSIPYVWPNWTEYISTSVH 180  
 QY 181 HVLWIHCHFTVYLVPSCIPFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240  
 DB 181 HVLWIHCHFTVYLVPSCIPFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISFATL 240  
 QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFLYCFISKRFPTMAAATL 300  
 DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFLYCFISKRFPTMAAATL 300  
 QY 301 KAFFKCKQKQPVQYTNHNFSSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350  
 DB 301 KAFFKCKQKQPVQYTNHNFSSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350

RESULT 13  
 ABP98724

ABP98724 standard; protein; 353 AA.	Db	61	SSNYLLAALAAADILVLFVDFVDFLEDFILNMOMPOVPDKIIIEVLEFSSIHSTWTV	120
ABP98724;	Qy	121	PLTIDRYIAVCHPLKYHTVSPARTKRVISVYITCFLTSIPYVYWPNTWEDYISTSVH	180
27-JUN-2003 (first entry)	Db	121	PLTIDRYIAVCHPLKYHTVSPARTKRVISVYITCFLTSIPYVYWPNTWEDYISTSVH	180
Human AXOR-57 protein.	Qy	181	HVLWIHCFVTYLVPCSIFFILNSIIVYKLRKNSFRLRGYSTGKTTAILFTITSIFATL	240
Antibacterial; antifungal; antiviral; antiprotozoal; analgesic;	Db	181	HVLWIHCFVTYLVPCSIFFILNSIIVYKLRKNSFRLRGYSTGKTTAILFTITSIFATL	240
cytostatic; antidiabetic; anorectic; anabolic; antiasthmatic; antagonist;	Qy	241	WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALLNTAINFFLYCFISKFRFTMAATL	300
antiparkinsonian; cardiac; hypotensive; hypertensive; nephrotropic;	Db	241	WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALLNTAINFFLYCFISKFRFTMAATL	300
osteopathic; antianginal; antitumor; antiallergic; antimigraine; agonist;	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
antimetetic; nootropic; tranquilizer; neuroprotective; antidepressant;	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
vaccine; gene therapy; G-protein coupled receptor; receptor; infection;	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
pain; cancer; diabetes; obesity; anorexia; bulimia; asthma; stroke;	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
Parkinson's disease; osteoporosis; angina pectoris; ulcer; allergy;	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
vomiting; schizophrenia; depression; dementia; Huntington's disease.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
Homo sapiens.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
OS	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
XX	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
GB2365009-A.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
13-FEB-2002.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
11-APR-2001; 2001GB-00009019.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
11-APR-2001; 2001GB-00009019.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
(SMIK ) SMITHKLINE BEECHAM CORP.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
(SMIK ) SMITHKLINE BEECHAM PLC.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
Elshourbagy N, Gattu M, Michalovich D, Shabon U;	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
WPI; 2003-203569/20.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
N-PSDB; ACC44115.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
New G-protein coupled receptor, AXOR 57, for diagnosing and treating	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
diseases, such as, pain, cancer, diabetes, obesity, anorexia, asthma,	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
Parkinson's disease, hypotension, hypertension, urinary retention, and	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
osteoporosis.	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
Claim 1; Page 27-28; 32pp; English.	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
The invention relates to the isolation of a novel G-protein coupled (7TM)	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
receptor AXOR 57 (this sequence) or a DNA fragment encoding a polypeptide	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
having at least 95% identity with AXOR-57 protein. The protein has been	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
shown to have homology to the human chromosome 16 clone C16B-H1_036A2	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
(AC008785). The protein and polynucleotides encoding the protein are used	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
in screening for compounds that stimulate or inhibit the function or	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
level of the polypeptide or polynucleotides, such as, agonists and	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
antagonists. The protein, polynucleotides, and antibodies to the protein	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
are used in diagnostic kits, to diagnose a disease. Polynucleotides	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
encoding the protein are used for chromosome localization studies, or for	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
tissue expression studies. The protein and nucleic acids encoding the	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
protein are used in vaccines for treating diseases such as, bacterial,	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
fungi, protozoal, and viral infections, pain, cancer, diabetes, obesity,	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
anorexia, bulimia, asthma, Parkinson's disease, acute heart failure,	Qy	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
hypertension, urinary retention, osteoporosis, angina	Db	301	KAPFFKCKQKOPVQFYTNHNFSTSSPWISPAISHCIKMLVYQYDKNGKEIK	350
pectoris, myocardial infection, stroke, ulcer, allergy, benign prostatic	Qy	301	KAPFFKCKQKOPVQ	

Qy	121	PLTIDRYIAVCHPLKXHTVSPARTRKVI	SVYITCF	LSIPYYWPNWNTEDYISTSVH	180
Db	121	PLTIDRYIAVCHPLKXHTVSPARTRKVI	SVYITCF	LSIPYYWPNWNTEDYISTSVH	180
Qy	181	HVLWIHCFTVYLVPCSF	FFILNSII	VYVKRKSNFRLRGYSTGKTTAILPTTIS	240
Db	181	HVLWIHCFTVYLVPCSF	FFILNSII	VYVKRKSNFRLRGYSTGKTTAILPTTIS	240
Qy	241	WAPRIIMILYHLYGAPIQNRMLVHIMS	DIANML	ALLANTAINFELYCFISKRPTMAAATL	300
Db	241	WAPRIIMILYHLYGAPIQNRMLVHIMS	DIANML	ALLANTAINFELYCFISKRPTMAAATL	300
Qy	301	KAFFKCQKQVQFYTNHFS	TSSPWI	SPANSHCICKMLVYQDKNGKPKI	350
Db	301	KAFFKCQKQVQFYTNHFS	TSSPWI	SPANSHCICKMLVYQDKNGKPKI	350

RESULT 15

ABB82499  
ID ABB82499 standard: protein: 353 AA.

AC ABB82499;

DT 22-JAN-2003 (first entry)

Human TGR20 polypeptide.

G-protein coupled receptor; GPCR; antipsoriatic; antiinflammatory; antileptic; nootropic; neuroprotective; antianemic; antiulcer; human; antiparkinsonian; antileptic; TGR20; receptor.

OS Homo sapiens.

WO200277001-A2

PD 03-OCT-2002.

08-MAR-2002: 2002WO-US007171.

PR 09-MAR-2001: 2001US-00802803-

16-MAR-2001; 2001US-0276649P.  
PR

PA (TULA-) TULARIK INC.

PI Tian H., Zhao J., Chen J., Cutler G:

WPI: 2003-018881/01.

DR N-PSDB: ABV73364.

New G-protein coupled receptor polypeptides and polynucleotides useful for identifying compounds for treating a TGR-associated disorder, e.g. peoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's disease, anemia.

PS Claim 15: Page 63: 87pp: English.

The invention relates to G-protein coupled receptor (GPCR) polypeptides and encoding polynucleotides selected from TGR20, TGR35, TGR36, TGR183, TGR341, TGR211, TGR216 and TGR79. The polypeptides and nucleic acids are useful for identifying compounds for treating a TGR-associated disorder, such as psoriasis, inflammatory bowel disease, hyperlipidemia, Parkinson's disease, Huntington's disease, anemia, immune and blood disorders, ulcerative colitis, Crohn's disease or spleen enlargement. They are also useful for identifying cells such as kidney, liver, hypothalamus, colon, adipose, or spleen cells, for forensics and paternity determination, diagnosing diseases and examining signal transduction. The present sequence represents a human TGR20 polypeptide.

**Query Match**

94.4%; Score 1857; DB 6; Length 353;

Best Local Similarity 100.0%; Pred. No. 1.9e-197;

[illegible]

Search completed: February 8, 2005, 18:25:21

Job time : 167 secs

THIS PAGE DELIBERATELY LEFT BLANK







```
QY 141 TyrProAlaAArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 432 TACCAGCCGCGACCCGAAAGTCATTTGTAAGTGTATTTACATCATCTGCTTCTCGACGAGC 491
QY 161 IleProTyrTyrTrpTyrProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 492 ATCCCTATTACTGTTGGTGGCCCAACATCTGGACTGAGAGACTACATCAGACACTCTGTGCAT 551
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 552 CACGTCCTCATCTGCATCCACCTGCTTACCGCTTACCTGTGGTGGCCCTGCTCCATCTTCTTC 611
QY 201 IleLeuAsnSerIleIleValTyrIleLeuArgLysSerAsnPheArgLeuArgGly 220
Db 612 ATCTTGAACTCAATCATTTGTGTACAGCTCAGGAGGAGGCAATTTTCTGCTCCGTCGC 671
QY 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 672 TACTCCAGGGGAGACACCGCCCATCTTGTTCACCATTTACCTCCATCTTTCACACACTT 731
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 732 TGGGCCCCCGCATCATCATGATTTCTTTACCACCTCTATGGGGCGCCCATCCAGACCGC 791
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 792 TGGCTGTGTACATCATCTCGACATTCGCCAACATGCTAGCCCTTCTGAAACACAGCCATC 851
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 852 AACTTCTTCTCTACTCTCTCATCAGCAAGCGGTTCCGCCACCATGGCAGCGCCGACGCTC 911
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 912 AAGGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCATATACTTTTCC 971
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 972 ATACCAAGTAGCCCTCGATCTCGCCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1031
QY 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360
Db 1032 CAGTATGACAAAATGAAACCTATATAAAGTCGTAAATGACAGCAAAAGCTCCTACCCAG 1091
QY 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372
Db 1092 TTTGAGATGCCATTGGAGCTTGTGTCATCATCTCTG 1127
```

```
RESULT 3
LOCUS CQ867782 2117 bp DNA linear PAT 13-SEP-2004
DEFINITION Sequence 1 from Patent WO2004074841.
ACCESSION CQ867782
VERSION CQ867782.1 GI:51997901
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
```

```
REFERENCE
AUTHORS Murphy,A.J. and Croll-Kalish,S.
TITLE Kor31-like proteins and methods of modulating kor31-mediated activity
JOURNAL Patent: WO 2004074841-A 1 02-SEP-2004;
REGENERON PHARMACEUTICALS, INC. (US); Murphy, Andrew J. (US);
Croll-Kalish, Susan (US)
FEATURES
source Location/Qualifiers
1..2117
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

ORIGIN

```
Alignment Scores: 5.14e-154 Length: 2117
Pred. No.: 1907.00 Matches: 372
Score: 83.22% Conservative: 0
Percent Similarity: 83.22% Mismatches: 0
Best Local Similarity: 96.95% Indels: 75
Query Match: 6 Gaps: 1
DB: 1
US-10-689-832-20 (1-372) x CQ867782 (1-2117)
QY 1 MetGluHisThrHisAlaHisIleLeuAlaAlaAsnSerSerIleSerTrpTrpSerProGly 20
Db 280 ATGAGACACAGCAGCCACCCACCTCGCAGCCACAGCTCGCTGCTTGGTGGTCCCCCGGC 339
QY 21 SerAlaCysGlyLeuGlyPheValProValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 340 TCGGCCCTGGCGCTTGGGTTTCGTGCCCGTGTCTACTACAGCTCTTGTGTGCTCGGT 399
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60
Db 400 TTACCAGCAAAATATCTTGACAGTGATCATCTCCAGCTGGTGGCAAGACAGAGAAG 459
QY 61 SerSerTyrAsnTyrIleLeuAlaLeuAlaAlaAspIleLeuValLeuPheIle 80
Db 460 TCTCTCTACACACTATCTTGGCAGCTCGCTGCTGCCGACATCTTGTGCTCTCTTTTCATA 519
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 520 GTGTTTGTGAGCTTCTGTGTGGAAGATTTTCATCTTGAACATGACAGATGCTCAGGTCCCC 579
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 580 GACAAGATCATAGAAGTGTGGAATTTCTCATCTCCATCCACACCTCCATATGGATTACTGTA 639
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 640 CCGTTAAACATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAAGTACCACACGGTCTCA 699
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 700 TACCAGCCCGCCAGCCGGAAGTCATTTGAAGTGTTCATCACCTGCTTCTTGACACAGC 759
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 760 ATCCCTATTACTTGTGGTGGCCCAACATCTGAGCTGAAGACTACATCAGACCTCTGTGCAT 819
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 820 CACGTCTCATCTGGATCCACTGCTTACCGTCTACCGTCTACCTGCTGCTCTCTTCTTTC 879
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
Db 880 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGAGAGCAATTTTCTGCTCCGTGGC 939
QY 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 940 TACTCCAGGGGAGACACCCGCTTGTTCACCATTTACCTTCATCTTTCACACACTT 999
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 1000 TGGGCCCCCGCATCATCATGATTTCTTTACCACCTCTATGGGGCGCCCATCCAGAACCCG 1059
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 1060 TGGCTGGTGCACATCATGTCCGACATTTGCCAACATGCTAGTCCCTTCTTGAAACAGCAGC 1119
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 1120 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGCCACCATGCGACCCCGCCGCTC 1179
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
```

Db 1180 AAGGCTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 1239  
Qy 321 IleThrSerSerProTrrpIleSerProAlaAsnSerHisCysIleIysMetLeuValTyr 340  
Db 1240 ATAACAAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGATC 1299

Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350  
Db 1300 CAGTATGACAAAATGGAAAACCTATAAAGTATCCCCGTGATTCCATAGGTGTGGCAAC 1359  
Qy 350 ----- 350

Db 1360 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGTGTCCCATCTATAGCTGAGCAGC 1419  
Qy 350 ----- 350

Db 1420 TCTCCTTAAGAGTGCTAATCCGATTTCTGTCTCCCGCAGACTGGGCAATTCAGACTG 1479  
Qy 350 ----- 350

Db 1480 GTAGATGAGAAGATGGGAAGAGAAGAGAGATGAAGCTGTGTTTACTTATGCA 1539  
Qy 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366

Db 1540 TTTATTTCCACAGAGTCGTGAATGACAGCAAAAGCTCTACCAAGTTTGAAGATGCCATTGG 1599  
Qy 366 yAlaCysValIleIleLeu 372  
Db 1600 AGCTTGTGTCAATCCTG 1618

RESULT 4  
AX254977 AX254977 1343 bp DNA linear PAT 10-OCT-2001  
LOCUS Sequence 21 from Patent WO0170978.  
DEFINITION AX254977  
ACCESSION AX254977  
VERSION AX254977.1 GI:16074504

KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Taupier, R. J., Majumder, K., Spaderna, S. K., Smithson, G., Mezes, P. S.  
and Vernet, C. A.

TITLE Polypeptides and nucleic acids encoding same  
JOURNAL Patent: WO 0170978-A 21 27-SEP-2001;  
Curagen Corporation (US)

FEATURES  
source  
1. .1343  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Alignment Scores:  
Pred. No.: 6,616-154 Length: 1343  
Score: 1903.00 Matches: 371  
Percent Similarity: 83.00% Conservative: 0  
Best Local Similarity: 83.00% Mismatches: 1  
Query Match: 96.75% Indels: 75  
DB: 6 Gaps: 1

US-10-689-832-20 (1-372) x AX254977 (1-1343)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
Db 2 ATGAGCACAGCACGCCACCTCGCAGCAACAGCTCGCTGTCTTGGTGTCCCGCGC 61

Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
Db 62 TCGGCTGGGGCTGGGGTTTCGTGGCCCGTGGTCTACTACAGCCCTCTGTGTGCTCGGT 121

Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgArgGlnLys 60

Db 122 TTACAGCAAAATATCTTGCAGAGTGCATCATCTCTCCAGCTGGTGGCAAGACAGAAG 181  
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80

Db 182 TCCTCTCAACAATATCTCTGGCACTCGCTCGCGACATCTTGGTCTCTTTTTCATA 241  
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100

Db 242 GTGTTTGGGACTTCTCTGTGGAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCCC 301  
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120

Db 302 GACAAGATCATAGAAGTGTCTGAAATTCATCTCCATCCACACCTCATATGATGATTACTGTA 361  
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140

Db 362 CCGTTAAACCATGACAGGTATATCATCTGTCTGCCACCGCTCAAGTACACACGGTCTCA 421  
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160

Db 422 TACCAGCCCGACCCCGAAGTCAATTGTAAGTGTTHACATCACTGCTTCTGACACAGC 481  
Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180

Db 482 ATCCCTATTACTGTGTGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 541  
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200

Db 542 CACGCTCTCATCTGGATCCACTGCTTACCGCTACCTGGTGGCTGCTCCATCTTCTTC 601  
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220

Db 602 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGGAGGAAGACAAATTTTCTCTCGTGGC 661  
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240

Db 662 TACTCCACGGGAGACCCACCGCATCTTGTTCACCATTTACCTCCATCTTGGCCACATT 721  
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260

Db 722 TGGGCCCCCGCATCATCATGATTTCTTACCACCTCTATGGGGCGCCCATCCAGAACCGC 781  
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280

Db 782 TGGCTGGGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAAACACACCATC 841  
Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300

Db 842 AACTTCTTCTCTACTGCTTCATCAGCAAGCGTTTCCGACCATGGCAGCGCCACGCTC 901  
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320

Db 902 AAGGCTTTCTTCAAGTGCAGAACACCTGTACAGTTCTACACCAATCATAACTTTTCC 961  
Qy 321 IleThrSerSerProTrrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340

Db 962 ATAAACAAGTAGCCCTCGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGTATC 1021  
Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350

Db 1022 CAGTATGACAAAATGGAAAACCTATAAAGTATCCCCGTGATTCCATAGGTGTGGCAAC 1081  
Qy 350 ----- 350

Db 1082 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGGTGCCCATCTATGGCTGAGCAGC 1141  
Qy 350 ----- 350

Db 1142 TCTCCTTAAGAGTGTAAATCCGATTTCTGTCTCCGCGAGACTGGGCAATTCACAGCTG 1201  
Qy 350 ----- 350





**Alignment Scores:**

Alignment scores:		
Pred. No.:	4,448-150	Length: 1062
Score:	1857.00	Matches: 350
Percent Similarity:	100.00%	Conservative: 0
Best Local Similarity:	100.00%	Mismatches: 0
Query Match:	94.41%	Indels: 0
DB:	6	Gaps: 0

US-10-689-832-20 (1-372) X BD183203 (1-1062)

Qy	1	MetGluHisThrHisAlaHisLeuAlaAlaHisSerSerLeuSerThrTrpSerProGly	20
Db	1	ATGAGCAGACGACGACGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGCGTGTGCTCCCGCGC	60
Qy	21	SerAlaCysAspLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly	40
Db	61	TCGGCTCGCGCTTGGGTTTGTGCGCCGTGGTGTACTACAGCCCTTGTGCTGTGCTCGGT	120
Qy	41	LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys	60
Db	121	TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAGAGACAGAG	180
Qy	61	SerSerTyrAsnTyrIleLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle	80
Db	181	TCCTCCTACAACTATCTCTTGGCACTCGCTGTCGCGACATCTTGGTCTCTTTTCA	240
Qy	81	ValPheValAspPheLeuLeuCluAspPheIleLeuAsnMetGlnMetProGlnValPro	100
Db	241	GTGTTTGTGGACTTCTGTGTGAAGAATTCATCTTGAAACATGTCAGATAGCCTCA	300
Qy	101	AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal	120
Db	301	GACAAGATCATAGAAGTGTGGAATCTCATCCATCCACACCTCCATATGGATTACTG	360
Qy	121	ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer	140
Db	361	CCGTTAAACATTTACAGAGTATATCGCTGTCTGCCACCGCTCAAGTACCAACAGCG	420
Qy	141	TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer	160
Db	421	TACCCAGCCCGCACCCGGAAGTCAATGTAAAGTGTTTACATCACCTGCTTCTTGAC	480
Qy	161	IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis	180
Db	481	ATCCCCATTACTGGTGGCCCAACATCTGGACTGAGACTACATCAGCACCTCTGTGCAT	540
Qy	181	HisValLeuIleTrpIleHisCysPheThrValTyrIleValProCysSerIlePhePhe	200
Db	541	CAGGTCTCATCTGGATCCACTGCTTTCACCGCTTACCTGGTGGCCCTGCTCCATCTT	600
Qy	201	IleLeuAsnSerIleValTyrLysIleuArgLysSerAsnPheArgLeuArgGly	220
Db	601	ATCTTGACTCAATCATTTGTGACAACTCATGAGGAGAGACAAATTTTGTGCTCGGTGC	660
Qy	221	TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu	240
Db	661	TACTCCACGGGGAGACACCGCCATCTTGTTCACCATTACCTCCATCTTGGCCACACT	720
Qy	241	TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg	260
Db	721	TGGGCCCGCCGATCATCATGATCTTTACCACTCTATGGGGCGCCATTCAGAACCGC	780
Qy	261	TrpLeuValHisIleMetSerAspIleAlaAsnMetIleAlaLeuLeuAsnThrAlaIle	280
Db	781	TGGCTGGTACACATCATGTCGACAAATGCCAACATGCTAGCCCTTCTGAAACACAGC	840
Qy	281	AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu	300
Db	841	AACTTCTTCTCTACTGTCTTCATCAGCAAGCGGTTCGCACCATGGAGCGCGCACGCTC	900
Qy	301	LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer	320
Db	901	AAGGCTTTCTTCAAGTGGCAGAAGCACTGTACAGTTCTACACCAATCATCACTTTTC	960

Qy	321	IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr	340
Db	961	ATACAAAGTAGCCCTGGAYCTCGCGGAAACTCACACTGCATCAAGATGCTGGTGAC	1020
Qy	341	GlnTyrAspLysAsnGlyLysProIleLys	350
Db	1021	CAGTATGACAAAAATGGAAAACCTTATAAAA	1050

## RESULT 8

LOCUS	BD183204	1062 bp	DNA	linear	PAT 17-JUN-2003
DEFINITION	Novel G protein-coupled		receptor and	its	DNA.

ACCESSION	BD183204
VERSION	BD183204.1
KEYWORDS	GI:31875404
SOURCE	JP 2002345481-A/4.
ORGANISM	Homo sapiens (human)
	Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 REFERENCE 1 (bases 1 to 1062)

2. Toebes L, de Boer EJ, Terao Y., Matsui, H. and Shintani, Y.  
Novel G protein-coupled receptor and its DNA  
Patent: JP 2002345481-A 4 03-DEC-2002;  
JOURNAL.

COMMENT	DN	JD	2002345481-2/4	OS	Homo sapiens (human)	TAKEDA CHEMICAL INDUSTRIES LTD	PATENT: JP 2002345481-A 4 03-DEC-2002;

PN JP 2002345481-A/4  
PD 03-DEC-2002  
PF 01-JUN-2001 JP 2001166688  
PF VASIKO TERAO HIDEKI MATSUI  
DT VASIKUSHI SUMITANI

PC A61P29/00, A61K45/00, A61P1/00, A61P3/00, A61P9/00, A61P25/00, PC A61P29/00,  
PC A61P25/00, A61P27/00, A61P14/70S, C08F16/78, C10M1/15, C10M1/19, BC

C12N1/21,  
PC A61F55/00,A61F57/00,C07K14/105,C07K16/28,C12N1/19, PC  
C12N1/21,  
PC C12N5/10,C12P21/02,C12Q1/68,G01N33/15,G01N33/50,G01N33/53, PC  
G01N33/53

GOIN33/33,  
PC GOIN33/566, C12N15/00, C12N5/00  
CC Novel G protein-coupled receptor and its DNA  
BU Rev. Location/Original

FH	key	Location/Qualifiers
PT	source	1..1062
PT		/organism='Homo sapiens (human)'
PT		Location/Qualifiers

```

FEATURES
  source
    location/qualifiers
      1. .1062
        /organism="Homo sapiens"
        /organism="Homo sapiens"

```

```
ORIGIN
/moi_type="genomic DNA"
/db_xref="taxon:9606"
```

Alignment Scores: 4.44e-150 Length: 1062  
Pred. No.: 2

Score:	1857.00	Matches:	350
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0

Query Match:	94.41%	Indels:	0
DB:	6	Gaps:	0
DB:			

US-10-689-832-20 (1-372) x BD183204 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaHisSerLeuSerTrpTrpSerProclly 20

Db  
1 ATGGAGCACACGACAGGCCACCTCGCAGCCACAGCTCGCTGTCTTGGTGTCTCCCCCGGC 60

Qy	21	SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly	40
Db	61	TCGGCCTCGCGGCTGGGTTCGTGCCGTGGTCTACTACAGCCTCTTGTCTGCCTCGGT	120

QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerClnLeuValAlaArgArgGlnIlys 60

DB	121	TTACCA	GAGAAATATCTTGACAGT	GATCATCTCTCCAGCTGGTGGCAAGACAGAAG	180
QY	61	SerSer	TyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle	80	

Db  
181 TCCTCTACAACTATCTCTTGGCACTCGCTGCTGCCACATCTTGGTCCTCTTTTCATA 240

FEATURES		Location/Qualifiers
source		1. 1062
		/organism="synthetic construct"
		/mol_type="unassigned DNA"
		/db_xref="taxon:32630"
		/note="Novel Sequence"
ORIGIN		
Alignment Scores:		4.44e-150 Length: 1062
Pred. No.:		1857.00 Matches: 350
Score:		100.00% Conservative: 0
Percent Similarity:		100.00% Mismatches: 0
Best Local Similarity:		94.41% Indels: 0
Query Match:		6 Gaps: 0
Db:		
US-10-689-832-20 (1-372) x AX498192 (1-1062)		
QY	1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGly	20
Db	1 ATGGAGCACAGCAGCCCACTCGCAGCCCAACAGCTCGCTGCTTGGTGGTCCCCGGC	60
QY	21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly	40
Db	61 TCGGCTCGGCTTGGGTTTCGTGCCGTGCTACTACAGCCTCTTGTGTGCTCGT	120
QY	41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys	60
Db	121 TTACCAGCAATAATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG	180
QY	61 SerSerTyrAsnTyrLeuLeuAlaAlaAspIleLeuValLeuPhePheIle	80
Db	181 TCCTCTACACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATA	240
QY	81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro	100
Db	241 GGTGTTGTGGACTTCTCTGTGGAGATTTCTTGAACATGCAGATGCCTCAGGTCCCC	300
QY	101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal	120
Db	301 GACAAGATCATAGAGTGTGGAATTTCTATCCATCCACACTCCATATGGAATTAATGTA	360
QY	121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer	140
Db	361 CGGTAAACATTCAGCAGTATATCGCTGTCTGCCACCCGCTCAAGTACACACGGTCTCA	420
QY	141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer	160
Db	421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTTACATCACCTGCTTCTCGACCAGC	480
QY	161 IleProTyrTyrTyrTyrProAsnIleTyrPheGluAspTyrIleSerThrSerValHis	180
Db	481 ATCCCTATTACTTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT	540
QY	181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe	200
Db	541 CACGTCCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGCCTGCTCCATCTTCTTC	600
QY	201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly	220
Db	601 ATCTTGAATCAATCATTTGTGTACAGCTCAGGAGGAAGACAAATTTTCGTCTCCGTGGC	660
QY	221 TyrSerThrGlySerThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu	240
Db	661 TACTCCAGGGAGACCAACCGCATCTTGTTCACCATTACCTCCATCTTGGCACACTT	720
QY	241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg	260
Db	721 TGGGCCCCCGCATCATCATGATTTCTTACACCTCTATGGGGCGCCCATCCAGAACCGC	780
QY	261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle	280
Db	781 TGGCTGGTGCAATCATGATGCCACATGTCACATGCTAGCCCTTCTGAAACACAGCCATC	840
QY	281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu	300
Db	841 AACTTCTTCTCTACTGCTTTCATCAGCAGCGGTTCCGACCATGGCAGCCCGCACGCTC	900
QY	301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer	320
Db	901 AAGGCTTTCTTCAAGTGCCAGGAAGCAACCTGTACAGTTCTACACCAATCATAACTTTTC	960
QY	321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr	340
Db	961 ATAACAAGTAGCCCTTGGATCTGCGCGCAAACTCACACTGCATCAAGATGTGGTGATC	1020
QY	341 GlnTyrAspLysAsnGlyLysProIleLys 350	
Db	1021 CAGTATGACAAAATGGAAAACCTATATAA 1050	
RESULT 9		
AX498192		1062 bp DNA linear PAT 26-SBP-2002
DEFINITION		Sequence 15 from Patent WO0242461.
ACCESSION		AX498192
VERSION		AX498192.1 GI:23343117
KEYWORDS		synthetic construct
SOURCE		synthetic construct
ORGANISM		other sequences; artificial sequences.
REFERENCE		1
AUTHORS		Chen, R., Chu, Z. L., Dang, H. T., Lowitz, K. P. and Pride, C.
TITLE		Endogenous and non-endogenous versions of human g protein-coupled receptors
JOURNAL		Patent: WO 0242461-A 15 30-MAY-2002; Arena Pharmaceuticals, Inc. (US)

Qy 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
 Db 841 AACTTCTCTCTACTCTTTCATCAGCAAGCGGTTCGCCACCATGGCGCGCCGCTC 900  
 Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
 Db 901 AAGCTTTCTTCAAGTCCGAGAACCACTGTACAGTTCTACACCAATCATAACTTTTCC 960  
 Qy 321 IleThrSerSerProTrrPileSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 Db 961 ATAACAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGATC 1020  
 Qy 341 GlnTyrAspLysAsnGlyLysProIleLys 350  
 Db 1021 CAGTATGACAAATAATGGAACCTATAAAA 1050

## RESULT 10

BD103851 1062 bp DNA linear PAT 27-AUG-2002  
 LOCUS Novel G protein-coupled receptor and its DNA.

## DEFINITION

BD103851

## ACCESSION

BD103851.1 GI:22649425

## VERSION

WO 0194582-A/3.

## KEYWORDS

Homo sapiens (human)

## SOURCE

Homo sapiens

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

## REFERENCE

1 (bases 1 to 1062)

## AUTHORS

Terao, Y., Matsui, H. and Shintani, Y.

## TITLE

Novel G protein-coupled receptor and its DNA

## JOURNAL

Patent: WO 0194582-A 3 13-DEC-2001;

TAKEDA CHEMICAL INDUSTRIES LTD, YASUKO TERAO, HIDEKI MATSUI, YASUSHI

SHINTANI

## COMMENT

OS Homo sapiens (human)

PN WO 0194582-A/3

PD 13-DEC-2001

PF 01-JUN-2001 WO 2001JP004643

PP 02-JUN-2000 JP OOP 170446, 23-JUN-2000 JP OOP 194926 PI

YASUKO TERAO, HIDEKI MATSUI, YASUSHI SHINTANI

PC C12N15/12, C12P21/02, C07K14/705, C07K16/28, A61K45/00, A61P25/00,

PC A61P29/00,

PC A61P9/00, A61P35/00, A61P3/00, A61P1/00, G01N33/566, PC

G01N33/50,

PC G01N33/15

CC Novel G protein-coupled receptor and its DNA

FH Key Location/Qualifiers

FT source 1..1062

FT /organism='Homo sapiens (human)'.  
 Location/Qualifiers

1..1062

/organism='Homo sapiens'

/mol\_type='genomic DNA'

/db\_xref='taxon:9606'

## ORIGIN

Alignment Scores:

Score: 4.44e-150 Length: 1062

Matches: 350

Conservative: 0

Mismatches: 0

Indels: 0

Gaps: 0

DB: 6

US-10-689-832-20 (1-372) x BD103851 (1-1062)

Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAhrArgGlnLys 60  
 Db 121 TTACCAAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 180  
 Qy 61 SerSerTyrAsnTyrIleLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
 Db 181 TCTCTCTACAACTATCTCTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTTCATA 240  
 Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
 Db 241 GTGTTTGTGGACTTCTCTGTGGAGATTTCATCTTGAACATGACGATGCTCTAGGTCCC 300  
 Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120  
 Db 301 GACAAGATCATAGAAGTCTGGAATTCTCATCATCCACACCTCCATATGGATTACTGTA 360  
 Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 Db 361 CCGTTAAACCAATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACCAACACGGTCTCA 420  
 Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 Db 421 TACCCAGCCCGCACCCCGGAAAGTCATTTAGTGTATTACATCACTGCTCTCTGACACG 480  
 Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
 Db 481 ATCCCTTACTTACTGGTGGCCCAACATCTGAGCTGAAGACTTACATCAGCACCTCTGTGCAT 540  
 Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 Db 541 CAGCTCTCATCTGGATCCACTCTTTCACGCTTACCTGCTGCTGCTCTCATCTCTTCTTC 600  
 Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220  
 Db 601 ATCTTGAACCTCAATCATCTGTGTACAGCTCAGAGGAGAGCAATTTTCTGCTCCGTCGC 660  
 Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
 Db 661 TACTCCACGGGAGACCCAGCCCATCTTGTTCACCACTTACCTCCATCTTTGGCCACACTT 720  
 Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
 Db 721 TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
 Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280  
 Db 781 TGGCTGGTACACATCATGTCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCCATC 840  
 Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
 Db 841 AACTTCTCTCTACTGCTTTCATCAGCAAGCGGTTCGACACCATGGCGCGCCGCGCTC 900  
 Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
 Db 901 AAGGCTTTCTTCAAGTCCAGAACCACTGTACAGTTCTTACACCAATCATAACTTTTCC 960  
 Qy 321 IleThrSerSerProTrrPileSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 Db 961 ATAACAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGATC 1020  
 Qy 341 GlnTyrAspLysAsnGlyLysProIleLys 350  
 Db 1021 CAGTATGACAAATAATGGAACCTATAAAA 1050

## RESULT 11

BD103852

## LOCUS

Novel G protein-coupled receptor and its DNA.

BD103852

## DEFINITION

BD103852.1 GI:22649426

## ACCESSION

WO 0194582-A/4.

## KEYWORDS

Homo sapiens (human)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20

Db 1 ATGAGCACACGACGCGCCACTCGACGCAACAGCTCGTGTCTTGGTGGTCCCGGCG 60

Qy 21 SerAlaCysGlyLeuGlyPheValProValTyrTyrSerLeuLeuLysLeuGly 40

Db 61 TCGGCTGCGGCTTGGGTTTTCGTCGCGCTGCTTACTACAGCTCTTGTGTCGCTCGGT 120



```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1062)
AUTHORS Terso,Y., Matsui,H. and Shintani,Y.
TITLE Novel G protein-coupled receptor and its DNA
JOURNAL Patent: WO 0194582-A 4 13-DEC-2001;
TAKEDA CHEMICAL INDUSTRIES LTD,YASUKO TERAU,HIDEKI MATSUI, YASUSHI
SHINTANI
COMMENT OS Homo sapiens (human)
PN WO 0194582-A/4
PD 13-DEC-2001
PF 01-JUN-2001 WO 2001JP004643
PR 02-JUN-2000 JP OP 170446,23-JUN-2000 JP OP 194926 PI
YASUKO TERAU,HIDEKI MATSUI,YASUSHI SHINTANI
PC C12N15/12,C12P21/02,C07K14/705,C07K16/28,A61K45/00,A61P25/00,
A61P29/00,
PC A61P9/00,A61P35/00,A61P37/00,A61P1/00,G01N33/566, PC
G01N33/50,
PC G01N33/15
CC Novel G protein-coupled receptor and its DNA
FH Key Location/Qualifiers
FT source 1..1062
FT /organism='Homo sapiens (human)'.
FEATURES
Source
Location/Qualifiers
1..1062
/organism='Homo sapiens'
/mol_type='genomic DNA'
/db_xref='taxon:9606'
ORIGIN
Alignment Scores:
Pred. No.: 4,44e-150 Length: 1062
Score: 1857.00 Matches: 350
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 94.41% Indels: 0
DB: 6 Gaps: 0
US-10-689-832-20 (1-372) x BD103852 (1-1062)
QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
Db 1 ATGGAGCACAGCAGCCACCTCGCAGCCACAGCTCGCTGTCTGTGGTGGTCCCGGC 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTrpSerLeuLeuLeuCysLeuGly 40
Db 61 TCGGCCTCGCGGTGGGTTCGTGGCGGTGGTCTACTACAGGCTCTTGTGTGCTCGGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60
Db 121 TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAGAAG 180
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 181 TCCTCTACACTATCTCTGGCACTCGCTGTGGCGACATCTTGTGCTCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 241 GTGTTGTGGACTTCCTGTTGGAAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCC 300
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrIleTrpIleThrVal 120
Db 301 GACAGATCATGAAGTGTGGAATTTCTCATCTCCACCTCCCATATGGAATTAATGTA 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 361 CCGTTAACCATTTGACAGGATATATCGCTGTCTGCCACCCCTCAAGTACCACAGGTCTCA 420
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 421 TACCCAGCCCGCACCCGGAAGTCAATGTTAAGTGTTTACATCACCTGCTCTCTGACACG 480

```

```

QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 481 ATCCCTTATTAATCTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 541 CACGTCCTCATCTGATCCACTGCTTACCCGCTTACCTGGTGGTCCCTGCTCCATCTCTTC 600
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 601 ATCTTGAATCAATCATTTGTGTACAGCTCAGAGAGAGACAAATTTTCTCTCCGTGTC 660
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 661 TACTCCACGGGGAAGACCCACCCCATCTTGTTCACCATTAACCTCTTTGCCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 721 TGGGCCCCCGCATCATGATCTTTTACCACCTCTATGGGGCGCCCATCCAGAACC 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
Db 781 TGGCTGGTGCACATCATGTCCGACATTCGAACATGCTAGCCCTTCTGAACACAGCCATC 840
QY 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 841 AACTTCTTCTTCTACTGCTTCATCAGCAAGCGTTCCGACCATGGCAGCCGCGCTC 900
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 901 AAGGCTTTCTTCAAGTGCAGAAAGCACTGTACAGTTCTACACCAATCATTAACCTTTC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 961 ATAAACAAGTAGCCCTCGATCTCGCGCGCAAACTCACACTGCATCAAGATGTGGTGATAC 1020
QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
Db 1021 CAGTATGACAAAATGGAACCTATATAA 1050
RESULT 12
AY635179 1167 bp mRNA linear PRI 04-JUL-2004
LOCUS Homo sapiens G protein-coupled receptor 139 (GPR139) mRNA, complete
DEFINITION cds:
ACCESSION AY635179
VERSION AY635179.1 GI:49413551
KEYWORDS .
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Complete coding sequence of GPR139
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1167)
AUTHORS Bonner,T.I., Nagle,J.W. and Kauffman,D.
TITLE Direct Submission
JOURNAL Submitted (25-MAY-2004) Lab of Genetics, NIMH, Bldg 36, Rm 3D06,
MSC4094, Bethesda, MD 20892-4094, USA
FEATURES
source
1..1167
/organism='Homo sapiens'
/mol_type='mRNA'
/db_xref='taxon:9606'
/chromosome='16'
/map='16p12.3'
/tissue_type='brain'
1..1167
/gene='GPR139'
92..1153
/gene='GPR139'

```



```
/codon_start=1
/product="G protein-coupled receptor 139"
/protein_id="AAT65818.1"
/db_xref="GI:49413552"
/tranlations="MEHTAHLAANSSLSWSPGACGLGFVVVYVYLLLCIGLPAN
ILTVILSOLVABROKSSNYLALAAADILVLFVDFVDFLEEDFLANQMOPQVDPK
IIVLFSSHTSIWITVPLIDRYIAVCHPLKHTVSPARTKRVISVITCFPLTS
IPYWPNIWTDIISTVHVLIIHCFVTVLVPCSIIFILMSIIVKLRKSNFRL
RGYTKTALITFTISFATLWAPRIIMLYHLYGAPIQNRWLHIMSDIANMLALL
NTAINFLCFISKRPRTRMAAATLKAFFKQKQPVQFYTNHNFISITSPWISPAHSC
IKMLVQYDNGKPKIKVSP"
primer_bind complement (1142..1167)
ORIGIN
US-10-689-832-20 (1-372) x AY635179 (1-1167)
Qy 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerSerLeuSerTrpTrpSerProGly 20
Db 92 ATGAGGACACAGCGACGCCACCTCGCAGCAACAGCTCGCTGCTTGGTGGTCCCGCGC 151
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
Db 152 TCGGCTTCGCGCTGGGTGTTGGTGGCGGTGCTTACTACAGCTCTTGGTGGTCCGCT 211
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAArgGlnLys 60
Db 212 TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 271
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAAspIleLeuValLeuPhePheIle 80
Db 272 TCCTCTCTACAACTATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA 331
Qy 81 ValPheValAspPheLeuLeuGluAAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 332 GTGTTTGTGGACTTCCTGTTGGAGATTTTCATCTTGAACTAGCAGATGCTCAGGTCCCC 391
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120
Db 392 GACAGATCATAGAGTCTGGATCTTCATCATCATCCACCTCCATGATGATCTGTA 451
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 452 CCGTTAACCATTTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 511
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 512 TACCCAGCCCGCACCCCGAAGTCATTGTAAGTGTTTTACATCATCCTGCTTCTTCGACGAC 571
Qy 161 IleProTyrTyrTrpTrpProAsnIleThrThrGluAspTyrIleSerThrSerValHis 180
Db 572 ATCCCTATTACTGTGGGCCCAACATCTGACTGAGACTATCATGACACTCTGTGAT 631
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 632 CAGCTCTCATCTGGATCCACTGCTTCACCGTCTACCTGTGCTGCTGCTCTCATCTCTTCT 691
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 692 ATCTTGAACATCAATATTGTATGTAAGCTCAGGAGGAAGCAATTTTCGTCTCCGTGGC 751
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 752 TACTCCAGGGGAGACACCGCACTTGTGTACCAATTAACCTTCCATCTTTGCAACTT 811
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
```

```
812 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCCG 871
261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAenThrAlaIle 280
872 TGGTGGTTCATCATCATGTCGCAATGTCACATGCTAGCCCTTCTGACACAGCCATC 931
281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
932 AACTTCTCTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATGTCGCGCAGCGC 991
301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
992 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTCTTACACCAATCATAACTTTTC 1051
321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
1052 ATACCAAGTAGTCCCTTGGATCTCGCGGCAACTACATGTCATGATGCTGGTGTAC 1111
341 GlnTyrAspLysAsnGlyLysProIleLys 350
1112 CAGTATGACAAAATGCAAAACCTATAAAA 1141
RESULT 13
AX664703 1062 bp DNA linear PAT 22-MAR-2003
LOCUS Sequence 12 from Patent WO02074960.
DEFINITION AX664703
ACCESSION AX664703.1 GI:29164463
VERSION
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Leiby, K.R., Kapeller-Libermann, R. and Glucksmann, M.
TITLE 38650, 28472, 5495, 65507, 81588 and 14354 methods and compositions
JOURNAL of human proteins and uses thereof
Patent: WO 02074960-A 12 SEP-2002;
Millennium Pharmaceuticals, Inc. (US)
FEATURES
source 1. .1062
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN
Alignment Scores: 1.12e-147 Length: 1062
Pred. No.: Matches: 344
Score: 1829.00
Percent Similarity: 98.86% Conservative: 2
Best Local Similarity: 98.29% Mismatches: 4
Query Match: 92.98% Indels: 0
DB: Gaps: 6
US-10-689-832-20 (1-372) x AX664703 (1-1062)
Qy 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerSerLeuSerTrpTrpSerProGly 20
Db 1 ATGGAGCACACGACGCCACCTCGCAGCAACAGCTCGCTGCTTGGTGGTCCCGCGC 60
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
Db 61 TCGGCTTCGCGCTGGGTGTTGGTGGCGGTGCTTACTACAGCTCTTGGTGGTCCCGCGC 120
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAArgGlnLys 60
Db 121 TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTCGCGACAGACAGAG 180
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAAspIleLeuValLeuPheIle 80
Db 181 TCCTCTCTACAACTATCTTGGCACTGCTGCTGCGCACTCTTGGTCTCTTTTCATA 240
```

```
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 241 GTGTTTGTGGACTTCCTGTTGGAAGATTTTCATCTTGAACATGACAGATGCTCAGGTCCCC 300
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 301 GACAAGATCATAGAAGTGTGGAATTTCTCATCCATCCACACCTCCATATGGAATTACTGTA 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 361 CGTTAAACATGACAGGATATTCGTGTCTGCCACCCCTCAAGTACCACACGGTCTCA 420
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 421 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTCATACCTGCTTCTCGACCAGC 480
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 481 ATCCCTATTTACTGTGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 541 CAGTCTCATCTGGGTCCACCTCTTACCGTCTACAGGTGCGCTGCTCCATCTTCTTC 600
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 601 ATCTTGAATCAATCGTTGTGTACAGCTCAGGAGGAGCAATTTTCTCTCTGTGTC 660
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 661 TACTCCACGGGAAGACCAACCGCATCTTGTTCACCATTTACCTTCCATCTTTGCCATCTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValAlaProIleGlnAsnArg 260
Db 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280
Db 781 TGGCTGTACACATCATGTCGACATTTGCCACATGCTAGCCCTTCTTGAACACAGCATC 840
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCACCATGGCAGCCGCGCATC 900
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 901 AAGCTTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATTAACCTTTCC 960
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 961 ATAACAAGTAGCCCTTGGATCTGCCGCGCAAACTCACACTGCATCAAGATGCTGTGTATC 1020
QY 341 GlnTyrAspLysAsnGlyLysProIleLys 350
Db 1021 CAGTATGACAAAATGAAAAACCTATAAAA 1050

RESULT 14
AX463235 LOCUS
DEFINITION Sequence 9 from Patent WO0248358.
ACCESSION AX463235
VERSION AX463235.1 GI:21886203
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS Smolyar, A., Zhu, Z., Encinas, J., Watanabe, S. and Okigami, H.
TITLE Regulation of human chemokine-like receptor
JOURNAL Patent: WO 0248358-A 9 20-JUN-2002;
Bayer Aktiengesellschaft (DE)
FEATURES
Location/Qualifiers
```

```
source 1..1062
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Alignment Scores:
Pred. No.: 2,37e-139 Length: 1062
Score: 1732.00 Matches: 332
Percent Similarity: 90.74% Conservative: 1
Best Local Similarity: 90.46% Mismatches: 0
Query Match: 88.05% Indels: 34
DB: 6 Gaps: 2
US-10-689-832-20 (1-372) x AX463235 (1-1062)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
Db 1 ATGGAGCACACGACGACGCCACCTCGCAGCCAAACAGCTCGTGTCTTGGTGGTCCCGGC 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 61 TCGGCTCGGGCTTGGGTTTGTGCCGTGTACTAGCCTCTTGTGTGCTCGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAargGlnLys 60
Db 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80
Db 181 TCCTCTCAACATATCTCTTGGCACTCGTGTGCGCATCTTGGTCTCTCTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 241 GGTGTTGGACTTCTCTGTGGAAGATTTCACTTGAACATGCATGCTCAGTGCCTCC 300
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 301 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACACCTCCATATGGAATTACTGTA 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 361 CGGTTAAACATTTGACAGGTATATCGTGTCTGCCACCCGCTCAAGTACACACGCTCTCA 420
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 421 TACCCAGCCCGCACCC----- 435
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSer----- 178
Db 436 -----TATTACTGTGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGCCCGG 489
QY 179 -----ValHisHisValLeu 183
Db 490 AAAGTCATTTGTAAGTGTTTTACATCACCTGCTTCTGACCAGCATCTGTCATCAGCTCTC 549
QY 184 IleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePheIleLeuAsn 203
Db 550 ATCTGGATCCCATGCTTACCGCTGTACCTGTGGGCCCTGCTCCATCTTCTTCACTTGAAC 609
QY 204 SerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGlyTyrSerThr 223
Db 610 TCAATCATTTGTACAACTCAGGAGGAGAGCAATTTTCCGTCTCCGTGGCTACTCCAGC 669
QY 224 GlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTrpAlaPro 243
Db 670 GGGAGAACCCCGCCATCTTGTTCACCATTTACCTCATCTTTGCCACACCTTTGGGCCCC 729
QY 244 ArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArgTrpLeuVal 263
Db 730 CGCATCATCATGATTTTACCACCTCTATGGGGCGCCCATCCAGAACCGCTGGTGTGTA 789
QY 264 HisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhePhe 283
```

```
Db 790 CACATCATGTCACACATGTCACATGTCAGCCCTTCTGAAACAGCCATCAACTCTCTTC 849
Qy 284 LeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeuLysAlaPhe 303
Db 850 CTCTACTGCTTCATCAGCAAGCGGTCGACACATGGCAGCCGACGCTCAAGGCTTTC 909
Qy 304 PheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnHisPheSerIleThrSer 323
Db 910 TTCAGTGCCAGAAAGCAACCTGTACAGTCTACACCAATCATAACTTTTCCATAACAGT 969
Qy 324 SerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyrGlnTyrAsp 343
Db 970 AGCCCTTGGATCTCGCGGCAAACTCACATGTCATCAAGATGTGGTGTACCAAGTATGAC 1029
Qy 344 LysAsnGlyLysProIleLys 350
Db 1030 AAAATGGAAACCTATANA 1050

RESULT 15
AX463232
LOCUS AX463232 1826 bp DNA linear PAT 15-JUL-2002
DEFINITION Sequence 6 from Patent WO0248358.
ACCESSION AX463232
VERSION AX463232.1 GI:21886202
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 Smolyar,A., Zhu,Z., Encinas,J., Watanabe,S. and Okigami,H.
AUTHORS Regulation of human chemokine-like receptor
TITLE Patent: WO 0248358-A 6 20-JUN-2002;
JOURNAL Bayer Aktiengesellschaft (DE)
FEATURES
source
1. 1826
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
ORIGIN

Alignment Scores:
Pred. No.: 5,21e-134 Length: 1826
Score: 1673.00 Matches: 330
Percent Similarity: 81.48% Conservative: 0
Best Local Similarity: 81.48% Mismatches: 0
Query Match: 85.05% Indels: 75
DB: 6 Gaps: 1

US-10-689-832-20 (1-372) x AX463232 (1-1826)
Qy 43 AlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLysSerSer 62
Db 170 GCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAGTCTCTCC 229
Qy 63 TyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPheIleValPhe 82
Db 230 TACAACTATCTCTGGCACTCGTGTGCGACATCTTGTGCTCTCTTTTCATAGTGT 289
Qy 83 ValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValProAspLys 102
Db 290 GTGGACTTCTGTTGGAAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCGACAG 349
Qy 103 IleIleGluValLeuGluPheSerSerIleHisThrSerIleThrPheThrValProLeu 122
Db 350 ATCATAGAAGTGTGGAATTTCTATCCATCCACACCTCCATATGGATTAATGTCACGTTA 409
Qy 123 ThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSerTyrPro 142
Db 410 ACCATTTGACAGGTATATCGTGTCTGCCACCCCTCAAGTACCACACAGGCTCATACCCA 469
Qy 143 AlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSerIlePro 162
```

Search completed: February 12, 2005, 15:01:48  
Job time : 4341 secs

```
Db 470 GCCCGCACCCGAAAGTCATTGTAAGTGTATTACATCACCCTGCTTCTGACCAGCATCCCC 539
Qy 163 TyrTyrTyrTyrProAsnIleTyrThrGluAspTyrIleSerThrSerValHisIleVal 182
Db 530 TATTACTGTGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCATCAGCTC 589
Qy 183 LeuIleTyrPheIleHisCysPheThrValTyrLeuValProCysSerIlePhePheIleLeu 202
Db 590 CTCATCTGGATCCACCTGCTTACCGTCTACCTGGTGGCCCTGCTCCATCTTCTTCACTTC 649
Qy 203 AsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGlyTyrSer 222
Db 650 AACTCAATCATTTGTGTACAGCTCAGGAGGAGAGCAATTTTGTGCTCCGTGGCTACTCC 709
Qy 223 ThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeuTyrAla 242
Db 710 ACGGGAGACACCCGCCATCTTGTTCACCATTTACCTCTTTCGCCACACTTGGGCC 769
Qy 243 ProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArgTyrLeu 262
Db 770 CCGCGCATCATCATGATCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGCTGGCTG 829
Qy 263 ValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIleAsnPhe 282
Db 830 GTGCACATCATGTCCGACATTTGCCACATGCTAGCCCTTCTGAACACAGCCATCAACTTC 889
Qy 283 PheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeuLysAla 302
Db 890 TTCTCTACTGCTTTCATCAGCAAGCGTTCCGACCATGCGACGCCCGCCAGCTCAAGGCT 949
Qy 303 PhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnHisPheSerIleThr 322
Db 950 TTCTTCAAGTGCAGAAAGCAACTGTATACAGTTCTACACCAATCATATACTTTCCATAACA 1009
Qy 323 SerSerProTyrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyrGlnTyr 342
Db 1010 AGTAGCCCTTGGATCTCGCGGCAAACTCACATGTCATCAAGATGCTGGTGTACCAAGTAT 1069
Qy 343 AspLysAsnGlyLysProIle-Lys----- 350
Db 1070 GACAAAAATGGAAACCTATAAAGTATCCCGTGTATTCATAGTGTGGCAACTACTGTC 1129
Qy 350 ----- 350
Db 1130 CTCTGTCTAATCCATTTCCAGATGGGAAGGTGTCCATCTATGGCTGAGCAGCTCTCTCT 1189
Qy 350 ----- 350
Db 1190 TAAGAGTGTCTAATCCGATTTCTCTCCCGCAGACTGGGCAATTTCTCAGACTGTAGAT 1249
Qy 350 ----- 350
Db 1250 GAGAGAGATGGGAAGAGAGAGAGAGAGATGAAGCTGTTTTTACTTATGATGATTTATT 1309
Qy 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleGlyAlaC 368
Db 1310 TCCACAGACTCGTAATGACAGCAAAAGCTCTTACCAGTTTGAAGATGCCATTGGAGCTTG 1369
Qy 368 sValIleIleLeu 372
Db 1370 TGTCTATCATCTCTG 1382
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: February 12, 2005, 10:59:01 ; Search time 560 Seconds  
(without alignments)  
3932.400 Million cell updates/sec

Title: US-10-689-832-20  
Perfect score: 1967  
Sequence: 1 MEHTHAHLAANSLSWSPG.....NDSKSSYQFIDAIGACVIL 372

Scoring table: BLOSUM62 Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0  
Searched: 4390206 seqs, 2959870667 residues  
Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+ p2n.model -DRV=xlp  
-O=/cgn2\_1/USEPTO.spool.p/US10689832/runat\_08022005\_112209\_23229/app.query.fasta\_1.519  
-DB=N Geneseq\_16Dec04 -QPMT=fastap -SUFFIX=ring -MINMATCH=0 -1-LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPMT=ptto -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US10689832 @CGN 1 1 708 @runat\_08022005\_112209\_23229 -NCPU=6 -ICPU=3  
-NO MAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DSV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DLEXT=7

Database : N\_Geneseq\_16Dec04:.\*  
1: Geneseqn1980s:.\*  
2: Geneseqn1990s:.\*  
3: Geneseqn2000s:.\*  
4: Geneseqn2001as:.\*  
5: Geneseqn2001bs:.\*  
6: Geneseqn2002as:.\*  
7: Geneseqn2002bs:.\*  
8: Geneseqn2003as:.\*  
9: Geneseqn2003bs:.\*  
10: Geneseqn2003cs:.\*  
11: Geneseqn2003ds:.\*  
12: Geneseqn2004as:.\*  
13: Geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1967	100.0	1119	4 AAS15730	Aas15730 DNA encod
2	1967	100.0	1119	10 ADJ87766	Adj87766 G-coupled
3	1967	100.0	1119	12 ADI79322	Adi79322 NOV10a co
4	1967	100.0	1119	12 ADO56001	Ado56001 DNA encod
5	1967	100.0	1130	6 AAD29672	Aad29672 Human G-p

6	1967	100.0	2198	9 ACF05275	Acf05275 Human G-p
7	1939.5	98.6	1110	9 ACF05276	Acf05276 Human G-p
8	1939.5	98.6	2189	10 ADD18120	Add18120 Human G-p
9	1907	96.9	2117	13 ADR16433	Adr16433 Human KOR
10	1903	96.7	1343	4 AAS15731	Aas15731 DNA encod
11	1903	96.7	1343	10 ADJ87768	Adj87768 G-coupled
12	1903	96.7	1343	12 ADI79324	Adi79324 NOV11 cod
13	1903	96.7	1343	12 ADO56003	Ado56003 DNA encod
14	1903	96.7	2273	12 ADO28955	Ado28955 Human nov
15	1879	95.5	1526	6 AAD46858	Aad46858 Human 7TM
16	1879	95.5	1526	8 ACA60889	Acag60889 Human cdn
17	1879	95.5	1526	10 AAS57024	Abas7024 cDNA enco
18	1864	94.8	1776	10 ADF70587	Adf70587 Orphan re
19	1857	94.4	1059	6 ABN84269	Abn84269 Human che
20	1857	94.4	1062	6 ABK15562	Abk15562 cDNA enco
21	1857	94.4	1062	6 ABK15563	Abk15563 cDNA enco
22	1857	94.4	1062	6 ABT04873	Abt04873 Human G p
23	1857	94.4	1062	10 ACC44115	Acc44115 Human AXO
24	1857	94.4	1062	10 ACC71785	Acc71785 Human G p
25	1857	94.4	1202	8 ABV73364	Abv73364 Human TGR
26	1741	88.5	1038	8 ABV73373	Abv73373 Mouse TGR
27	1741	88.5	1038	12 ADO28957	Ado28957 Mouse nov
28	1732	88.1	1062	6 ABN84273	Abn84273 Human che
29	1673	85.1	1826	6 ABN84272	Abn84272 Human che
30	1628	82.8	957	12 ADI79347	Adi79347 NOV10b co
31	1628	82.8	957	12 ADO56026	Ado56026 DNA encod
32	1623	82.5	1002	6 AAD27501	Aad27501 Human G-p
33	1623	82.5	1002	10 AAD61658	Aad61658 Human GPC
34	1623	82.5	1032	6 ABN84271	Abn84271 Human che
35	1623	82.5	1070	6 ABN84270	Abn84270 Human che
36	1623	82.5	1158	6 ABZ42880	Abz42880 Human GPC
37	1623	82.5	1167	13 ADQ89945	Adq89945 Antagonis
38	1616	82.2	1000	6 AAD28102	Aad28102 Human thy
39	1607	81.7	946	10 ADJ87879	Adj87879 G-coupled
40	1471	74.8	864	4 AAS42811	Aas42811 Human G p
41	1447	73.6	867	6 AAS98134	Aas98134 Human DNA
42	1447	73.6	894	6 AAS98143	Aas98143 Human DNA
43	1431	72.2	963	10 ADC12695	Adc12695 Human GPC
44	1317	67.0	795	6 AAS98071	Aas98071 Human DNA
45	1317	67.0	795	8 ABZ42552	Abz42552 Human G p

ALIGNMENTS

RESULT 1  
AAS15730  
ID AAS15730 standard; cDNA; 1119 BP.  
XX  
AC AAS15730;

14-FEB-2002 (first entry)

DNA encoding chemokine receptor family related protein, NOV10.

NOV; cytostatic; psoriasis; nontropic; neuroprotectant;  
cerebroprotective; hepatic; antiinflammatory; diabetes;  
haemostatic; atherosclerosis; gene therapy; neurogenesis;  
differentiation; proliferation; haematopoiesis; wound healing;  
angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;  
haemophilia; allergy; Pendered syndrome; skeletal dysplasia;  
ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;  
chemokine receptor; chromosome 1; ss.

OS Homo sapiens.

Key	Location/Qualifiers
FT CDS	1..1119
FT	/*tag= a
FT	/product= "Chemokine receptor related protein, NOV10"
FT	1..141
FT	/*tag= b
FT	142..1116
FT	/*tag= c



Db 1021 CAGTATGACAAAATGAAACCTATATAAAGTCGTAATGACGAAAGCTCTACCAG 1080  
Qy 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
Db 1081 TTTGAAGATGCCATTGGAGCTTGTCATCATCTG 1116

RESULT 2

ADJ87766  
ID ADJ87766 standard; DNA; 1119 BP.

XX AC ADJ87766;

DT 06-MAY-2004 (first entry)

XX DE G-coupled protein receptor-related protein coding sequence #56.

XX KW novel protein; G-coupled protein receptor-related protein;  
KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;  
KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;  
KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolaemia;  
KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;  
KW viral infections; bacterial infection; parasitic infection;  
KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;  
KW Alzheimer's disease; tuberosus sclerosis; hypercalcaemia; cerebral palsy;  
KW gene; ds.

XX OS Unidentified.

XX PN W02002102321-A2.

XX PD 27-DEC-2002.

XX PF 18-JUN-2002; 2002WO-US019522.

XX PR 18-JUN-2001; 2001US-0298994P.

XX PR 04-OCT-2001; 2001US-0299134P.

XX PR 06-JUN-2002; 2002US-00972446.

XX PR 07-JUN-2002; 2002US-00298994.

XX PA (CURA-) CURAGEN CORP.

XX PI Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;

XX PI Pena CE, Spaderma SK, Zhong M;

XX DR WPI: 2003-167441/16.

XX DR P-PSDB; ADJ87767.

XX PT New MOLX polypeptides and polynucleotides, useful in gene therapy,

XX PT particularly for treating or preventing e.g. cardiomyopathy,

XX PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate

XX PT cancer, stroke or pain.

XX PS Claim 8; SEQ ID NO 201; 378pp; English.

XX CC The invention comprises the amino acid and coding sequences of novel G-  
CC coupled protein receptor-related (MOL) proteins. The DNA and protein  
CC sequences of the invention are useful for treating or preventing a MOL-  
CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders  
CC associated with cell signal processing and metabolic pathway modulation,  
CC or diabetes. The DNA and protein sequences are also useful for the  
CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,  
CC pain, hypercholesterolaemia, obesity, hypertension, Crohn's disease,  
CC systemic lupus erythematosus, viral infections, bacterial infections,  
CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau  
CC syndrome, Alzheimer's disease, tuberosus sclerosis, hypercalcaemia, or  
CC cerebral palsy. The present DNA sequence encodes a MOL protein of the  
CC invention.

XX SQ Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2.52e-187	Length:	1119
Score:	1967.00	Matches:	372
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0
US-10-689-832-20 (1-372) x ADJ87766 (1-1119)			
Qy	1	MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly	20
Db	1	ATGGAGCACACGACGCCACCTCGCAGCCACAGCTCGCTCTTGGTGGTCCCCCGGC	60
Qy	21	SerAlaCysGlyLeuGlyPheValProValValTyrSerLeuLeuLeuCysLeuGly	40
Db	61	TGGGCTCGGGCTTGGGTTTCGTGGCGGTGCTACTACAGCCTCTTGTGTGCTCGGT	120
Qy	41	LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys	60
Db	121	TTACCCAGCAATATCTTGACAGTGATCATCTCCAGCTGGTGGCAAGACAGAAG	180
Qy	61	SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPheIle	80
Db	181	TCTCTCTACACTATCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTCATA	240
Qy	81	ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro	100
Db	241	GTGTTTGTGGACTTCCTCTTGGAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCCC	300
Qy	101	AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal	120
Db	301	GACAGATCATAGAAAGTGTGGAAATTTCTCATCTCCACCTCCATATGGATTAATCTGA	360
Qy	121	ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer	140
Db	361	CGTTAACCTTGCACAGGTATATCGCTGCTGCGACCCCTCAAGTACACACGGTCTCA	420
Qy	141	TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer	160
Db	421	TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTATACATCACCTGCTTCTGTACACAGC	480
Qy	161	IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis	180
Db	481	ATCCCCATATTACTGTGGGCCCAACATCTGAGCTGAAGACTATCATCAGCAGCTCTGTCAT	540
Qy	181	HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe	200
Db	541	CACGTCCTCATCTGGATCCACTGCTTACCGTCTACCTGTGGCCCTGCTCCATCTTCTTC	600
Qy	201	IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly	220
Db	601	ATCTTGAACCTCAATCATTTGTGTACAGCTCAGGAGGAGAGCAATTTTCTCTCCGTGGC	660
Qy	221	TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu	240
Db	661	TACTCCAGCGGAGACACACCGCCATCTTGTTCACCATTACCTTCCATCTTTGACACATT	720
Qy	241	TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValaProIleGlnAsnArg	260
Db	721	TGGGCCCCCGCATCATCATGATTTTATACACCTCTATGCGGCGGCCCATCCAGAACCCG	780
Qy	261	TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle	280
Db	781	TGGCTGTACACATCATGTCCGACATTGCCACATGCTAGACCTTCTTGACACAGCCATC	840
Qy	281	AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu	300
Db	841	AACTTCTTCTCTACTGCTTTCATCAGCAGCGGTTCCGACCATCCAGCGCCGACGCTC	900
Qy	301	LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer	320
Db	901	AAGGCTTTCTTCAAGTGGCAGAAACCTGTACAGTTCTACACCAATCATAACTTTTTC	960

QY 321 IleThrSerSerProTrrPileSerProAlaAenSerHisCysIleLeuMetLeuValTyr 340  
 Db 961 ATAAACAGTAGCCCTGGATCTCGCGGCAAACTACACTGCATCAAGTGTGGTGTAC 1020  
 QY 341 GlnTyrAspLysAsnGlyLysProLeuLysSerArgAsnAspSerLysSerSerTyrGln 360  
 Db 1021 CAGTATGACAAAATGGAAACCTATAAAAGTCGTAAATGACAGCAAAAGCTCTTACCAG 1080  
 QY 361 PheGluAspAlaIleGlyAlaCysValIleLeu 372  
 Db 1081 TTGAAGATGCCATTGGAGCTTGTGTCAATCATCTG 1116  
 RESULT 3  
 ADI79322  
 ID ADI79322 standard; DNA; 1119 BP.  
 XX  
 AC ADI79322;  
 XX  
 XX 22-APR-2004 (first entry)  
 XX  
 XX NOV10a coding sequence, SEQ ID 19.  
 KW Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;  
 KW Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;  
 KW Ophthalmological; Antipsoriatic; Neuroprotective; Nootropic;  
 KW Antiartherosclerotic; Hypotensive; Cardiant; Cerebroprotective;  
 KW Gene Therapy; NOV; human; cancer; myelogenous leukemia;  
 KW congenital neonatal autoimmune thrombocytopenia; immunological disorder;  
 KW allergy; infection; asthma; lung disease; reproductive disorder;  
 KW haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;  
 KW diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;  
 KW hypertension; stroke; heart failure; chromosome 1; NOV10a;  
 KW chemokine receptor; gene; db.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 CDS 1..1119  
 FT /\*tag= a  
 FT /product= "NOV10a"  
 XX  
 XX WO2004009635-A2.  
 XX  
 XX 29-JAN-2004.  
 XX  
 XX 04-OCT-2001; 2001US-0081292.  
 XX  
 XX 20-MAR-2001; 2001US-00813432.  
 XX  
 XX (CURA-) CURAGEN CORP.  
 XX  
 PI Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;  
 PI Vernet CAM;  
 XX  
 XX WPI: 2004-123380/12.  
 DR P-PSDB; ADI79323.  
 XX  
 XX Claim 8; Page 41; 158pp; English.  
 XX  
 CC The present invention relates to novel NOVX proteins and their coding  
 CC sequences (ADI79304-ADI79327). The sequences are useful for the  
 CC manufacture of a medicament for treating a syndrome associated with a  
 CC human disease associated with the NOVX protein, or for diagnosing and treating  
 CC disorders associated with the NOVX protein, such as cancer, myelogenous  
 CC leukaemia, congenital neonatal autoimmune thrombocytopenia,  
 CC immunological disorders, allergy and infection, asthma, lung diseases,  
 CC reproductive disorders, male and female reproductive diseases,  
 CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,  
 CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,  
 CC hypertension, stroke and heart failure. NOV10a represents a new member of  
 CC the chemokine receptor family and the gene is located on chromosome 1.  
 XX  
 XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 2,528-187 Length: 1119  
 Score: 1967.00 Matches: 372  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 12 Gaps: 0  
 US-10-689-832-20 (1-372) x ADI79322 (1-1119)  
 QY 1 MetGluHisThrHisAlaHisLeuAlaAaenSerSerLeuSerTrrPrrSerProGly 20  
 Db 1 ATGGAGCACACGACGACGCCACCTCGACGCCAACAGCTCGTCTTGGTGGTCCCCGGC 60  
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
 Db 61 TCGGCTCGCGCTTGGGTTTGGTGGCGGTGTACTACAGCCTTGTGTGGCTCGGT 120  
 QY 41 LeuProAlaAenIleLeuThrValIleIleLeuSerGlnLeuValAlaAaenArgGlnLys 60  
 Db 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
 QY 61 SerSerTyrAenTyrLeuLeuAlaAlaAaenSerLeuLeuValLeuPhePheIle 80  
 Db 181 TCCTCTACAACTATCTTGGCACTCGTCTGCGACATCTTGGTCTCTTTTCATA 240  
 QY 81 ValPheValAppPheLeuLeuGluAspPheIleLeuAenMetGlnMetProGlnValPro 100  
 Db 241 GTGTTTGGGACTTCTCTGTTGGAAGATTTCATCTTGAACATGACAGTCCCTCAGGTCCC 300  
 QY 101 AspyIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrIleThrVal 120  
 Db 301 GACAAGATCATAGAGTGTGGAATTCATCTCCATCCACCTCCATATGATGATTACTGTA 360  
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 Db 361 CGGTAAACCATTGACAGGTATATCGTGTCTGCCACCGCTCAGTACACACAGGTCTCA 420  
 QY 141 TyrProAlaAaenThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 Db 421 TACCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTTACATCACCTGCTCTGTGACACG 480  
 QY 161 IleProTyrTyrTrrPrrProAsnIleThrPrrGluAspTyrIleSerThrSerValHis 180  
 Db 481 ATCCCTTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
 QY 181 HisValLeuIleTrrPrrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 Db 541 CAGGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGCTGCTCCATCTCTTTC 600  
 QY 201 IleLeuAenSerIleIleValTyrLysLeuArgLysSerAsnPhaArgLeuArgGly 220  
 Db 601 ATCTTGAACCTCAATCATTTGTGTACAAAGCTCAGGAGGAAGAGCAATTTTCTGTCTCG 660  
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
 Db 661 TACTCCAGGGGAAGACACCCGCCATCTTGTTCACCATTTACCTCATCTTTGCCACATT 720  
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValAlaProIleGlnAenArg 260  
 Db 721 TGGGCCCCCGCATCATCATGATCTTTTACCACTCTATGGGGCCGCCCATCCAGAACCGC 780  
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAenMetLeuAlaLeuLeuAenThrAlaIle 280  
 Db 781 TGGTGGTACACATCATCTCCGACATTCGCCAACATGTAGCCCTTCTCTGAACACAGCCATC 840  
 QY 281 AenPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
 Db 841 AACTTCTTCTCTACTGTCTTATCAGCAAGCGGTTCGACCATGTCAGCGCCCGCCAGCTC 900  
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAenPheSer 320



Db 901 AAGCTTTCTCAAGTCCAGAACCACTGTACAGTTCTACACCAATCAATACTTTTCC 960  
 Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 Db 961 ATAACAAGTAGCCCTGGATCTCGCCGGCAAACTCACACTGCATCAAGATGCTGGGTAC 1020  
 Qy 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360  
 Db 1021 CAGTATGACAAAAATGGAACCTATATAAAGTCGTAATGACAGCAAAAGCTCTACCAG 1080  
 Qy 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
 Db 1081 TTGGAGATGCCATTGGAGCTTGTGTCATCATCTCTG 1116

## RESULT 4

ADO56001

ID ADO56001 standard; cDNA; 1119 BP.

XX AC

XX ADO56001;

XX DT

XX 15-JUL-2004 (first entry)

XX DE DNA encoding human NOV10.

XX KW

XX human; gene; ss; cancer; obesity; diabetic nephropathy;

XX KW acute pancreatitis; stroke; multiple sclerosis.

XX OS

XX Homo sapiens.

XX PN

XX US2004058862-A1.

XX PD

XX 25-MAR-2004.

XX PF

XX 18-SEP-2002; 2002US-00246583.

XX PR

XX 18-SEP-2002; 2002US-00246583.

XX XX

XX PA (NAJU/) MAJUMDER K.

XX PI

XX Majumder K;

XX DR

XX WPI; 2004-268835/25.

XX P-PSDB; ADO56002.

XX XX

XX Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic nephropathy, acute pancreatitis, strokes and multiple sclerosis.

XX PT

XX Claim 9; Page 26-27; 87pp; English.

XX PS

XX The invention relates to novel isolated NOVX nucleic acids and encoded polypeptides. The nucleic acids, polypeptides and antibodies raised against the polypeptides are useful for preventing or treating diseases associated with aberrant NOVX expression or activity e.g., cancer, obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple sclerosis. The present sequence represents a NOVX nucleic acid of the invention.

XX SQ

XX Sequence 1119 BP; 258 A; 353 C; 220 G; 288 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	2.52e-187	Length:	1119
Score:	1967.00	Matches:	372
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	12	Gaps:	0

US-10-689-832-20 (1-372) x ADO56001 (1-1119)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20

Db 1 ATGGAGCACACGACGCCCACTCGCAGCCAAACAGCTCGCTGTCTTGGTGGTCCCGCGGC 60

RESULT 5

Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
 Db 61 TCGCCCTGCGGCTTGGGTTTCGTGCCCGGTCTACTACAGCCTCTTGTGTGCTCGGT 120  
 Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60  
 Db 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAGAAACAGAAG 180  
 Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
 Db 181 TCTCTCTACAATCTCTTGGCACTCGCTGGCGACATCTTGGTCTCTTTTTCATA 240  
 Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
 Db 241 GTGTTTGTGGACTTCTGTGTGAAGATTTTCATCTTGAACATGCAGATGCTCAGGTCCC 300  
 Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
 Db 301 GACAAGATCATAGAAGTGTGGAATTTCTCATCTCCACCTCCATATGGATTACTGTA 360  
 Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 Db 361 CGGTTAACCATTTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
 Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 Db 421 TACCCAGCCCGCACCGGAAAGTCATTGTAAGTGTTTTACATCACCTGTCTTCTGACGAC 480  
 Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
 Db 481 ATCCCCCTATTACTGGTGGCCCAACATCTGAGCTGGAAGACTACATCAGCACTCTGTGCAT 540  
 Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 Db 541 CAGCTCTCATCTGGATCCACTGCTTACCGTCTACTGTGGTGGCTGCTCCATCTTCTTC 600  
 Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
 Db 601 ATCTTGAATCAATCATTTGTGTACAAGCTCAGAGGAGAGAGCAATTTTCTGCTCCGTGC 660  
 Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
 Db 661 TACTCCACGGGAGAACCCACCGCCATCTTGTTCACCATTTACCTTCCATCTTTGCCACAT 720  
 Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValProIleGlnAsnArg 260  
 Db 721 TGGGCCCCCGGCATCATCATGATTTCTTACCACTCTATGGGGGCGCCATCCAGAACCGC 780  
 Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
 Db 781 TGGCTGGTACACATCATGTCCGACATTTGCCAATGTAGTCCCTTCTGAACACAGGCATC 840  
 Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
 Db 841 AACTTCTTCTCTACTGTCTCATCAGACGCGGTTCGCCACCATGGCAGCGCCGACGCTC 900  
 Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
 Db 901 AAGGCTTCTTCAAGTGGCCAGAACCACTGTACAGTTCTACACCAATCATAACTTTTCC 960  
 Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 Db 961 ATAACAAGTAGCCCTGGATCTCGCCGGCAAACTCACACTGCATCAAGATGCTGGGTAC 1020  
 Qy 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360  
 Db 1021 CAGTATGACAAAAATGGAACCTATATAAAGTCGTAATGACAGCAAAAGCTCTACCAG 1080  
 Qy 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
 Db 1081 TTGGAGATGCCATTGGAGCTTGTGTCATCATCTCTG 1116

AAD29672  
 ID AAD29672 standard; cDNA; 1130 BP.  
 XX  
 AC AAD29672;  
 XX  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human G-protein coupled receptor (GCRC-6) cDNA.  
 XX  
 KW Human; G-protein coupled receptor; GCRC-6; cell proliferative disorder;  
 KW neurological; cardiovascular; gastrointestinal; autoimmune; inflammatory;  
 KW metabolic; hepatitis; psoriasis; cancer; epilepsy; Alzheimer's disease;  
 KW Pick's disease; Huntington's disease; Parkinson's disease; hypertension;  
 KW atherosclerosis; myocardial infarction; gastritis; cirrhosis; cytostatic;  
 KW osteoporosis; Crohn's disease; acquired immunodeficiency syndrome; AIDS;  
 KW anaemia; asthma; rheumatoid arthritis; diabetes; obesity; drug screening;  
 KW transgenic animal; allergy; gene therapy; hepatotropic; anticonvulsant;  
 KW neurotropic; neuroprotective; cardiant; immunosuppressive; anorectic;  
 KW viricide; gene; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 12..1130  
 FT /\*tag= a  
 FT /product= "Human GCRC-6"  
 FT sig\_peptide 12..152  
 FT /\*tag= b  
 FT mat\_peptide 153..1127  
 FT /\*tag= c  
 FT /product= "Human mature GCRC-6"  
 XX  
 WO200210387-A2.  
 XX  
 PD 07-FEB-2002.  
 XX  
 XX 25-JUL-2001; 2001WO-US023433.  
 XX  
 XX 27-JUL-2000; 2000US-0221478P.  
 XX 03-AUG-2000; 2000US-0223268P.  
 XX 21-AUG-2000; 2000US-0227054P.  
 XX 08-SEP-2000; 2000US-0231121P.  
 XX 13-SEP-2000; 2000US-0232243P.  
 XX 15-SEP-2000; 2000US-0232691P.  
 XX 22-SEP-2000; 2000US-0235146P.  
 XX  
 XX (INCY-) INCYTE GENOMICS INC.  
 XX  
 XX Thornton M, Patterson C, Lal P, Burford N, Yue H, Gandhi AR;  
 XX Elliot VS, Ramkumar J, Baughn MR, Kallick DA, Wallia NK, Hafalia AJA;  
 XX Yao MG, Lu Y, Tribouley CM, Policky JL, Kearney L, Gaul RC;  
 XX Warren BA, Lee EA, Ding L;  
 XX  
 XX WPI; 2002-188744/24.  
 XX P-PSDB; AAE18645.  
 XX  
 XX New human G-protein coupled receptor polypeptide for diagnosis,  
 XX prevention and treatment of cell proliferative, neurological,  
 XX cardiovascular, gastrointestinal, autoimmune/inflammatory, and metabolic  
 XX disorders.  
 XX  
 XX Claim 5; Page 141; 150pp; English.  
 XX  
 XX The invention relates to novel human G-protein coupled receptors (GCRC)  
 XX and their encoding polynucleotides. GCRC is useful as an immunogen for  
 XX preparing monoclonal and polyclonal antibodies. GCRC is useful for  
 XX diagnosing, treating and preventing a cell proliferative disorder (e.g.,  
 XX hepatitis, psoriasis, cancer), a neurological disorder (e.g., epilepsy,  
 XX Alzheimer's disease, Pick's disease, Huntington's disease, Parkinson's  
 XX disease), a cardiovascular disorder (e.g., atherosclerosis, hypertension,  
 XX myocardial infarction), gastrointestinal disorder (e.g., gastritis,  
 XX cirrhosis, Crohn's disease), an autoimmune/ inflammatory disorder (e.g.,  
 XX acquired immunodeficiency syndrome (AIDS), allergy, anaemia, asthma,

CC rheumatoid arthritis), a metabolic disorder (e.g., diabetes, obesity,  
 CC osteoporosis), and viral infections. GCRC is useful in a number of drug  
 CC screening techniques, and to analyse the proteome of a tissue or cell  
 CC type. GCRC is useful for creating knockin humanised animals or  
 CC transgenic animals to model human diseases, in somatic or germline gene  
 CC therapy, to generate a transcript image of a tissue or cell type, for  
 CC detecting differences in the chromosomal location due to translocation,  
 CC inversion, etc., among normal, carrier or affected individuals, and as  
 CC hybridization probes for mapping naturally occurring genomic sequences.  
 CC GCRC is useful in Southern or northern analysis, dot blot or other  
 CC membrane-based technologies, in PCR technologies, in dipstick, pin,  
 CC multiplexed enzyme linked immunosorbent (ELISA)-like assays, and in  
 CC microarrays utilising fluids or tissues from patients to detect altered  
 CC GCRC expression. The present sequence is human GCRC-6 cDNA  
 XX  
 SQ Sequence 1130 BP; 258 A; 357 C; 226 G; 289 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 2,55e-187 Length: 1130  
 Score: 1967.00 Matches: 372  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 6 Gaps: 0  
 US-10-689-832-20 (1-372) x AAD29672 (1-1130)  
 QY 1 MetGluHisThrHisAlaHisLeuAlaAlaHisSerSerLeuSerTrpTrpSerProGly 20  
 DB 12 ATGGAGCACAGCAGCCGACCTCGCAGCCACAGCTCGTCTGTGGTGGTCCCGGC 71  
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40  
 DB 72 TCGGCTCGCGCTGGGTTCTGTCGCTGCTACTACGCTCTTCTGTGCTCGGT 131  
 QY 41 LeuProAlaAsnLeuLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60  
 DB 132 TTACAGCAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAGAGACAGAG 191  
 QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAlaAlaAlaAlaAlaAlaPhePhe 80  
 DB 192 TCCTCCTACAACTATCTCTGGCCTCGCTGTCGCGACATCTTGGTCTCTTTTCATA 251  
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleIleLeuMetGlnMetProGlnValPro 100  
 DB 252 GTGTTTGTGGACTTCTCTGTGGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCC 311  
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
 DB 312 GACAGATCATAGAGTCTGGATTCATCTCCATCCACCTCCATATGATGATCTGTA 371  
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 DB 372 CCGTTAACCATTCAGAGTATATCGCTGTCGCCACCGCTCAAGTACACACGCTCTCA 431  
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 DB 432 TACCAGCCCGACCCGGAAGTCAATTGTAAGTGTTCATACACCTGCTCTCTGACACG 491  
 QY 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
 DB 492 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGACCTCTGTGCAT 551  
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 DB 552 CACGCTCTCATCGATCCATCTGCTTACCGCTTACCTGGTGGCTGCTGCTCTCTCTTC 611  
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220  
 DB 612 ATCTTGAACATCAATGTTGTAGAGCTCAGGAGGAGAGCAATTTTCGCTCGGTGGC 671  
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240

Db 672 TACTCCAGGGGGAAGACACCGGCATCTTGTTCACCATTAACCTCCATCTTTGCCACACTT 731  
 Qy 241 TrpAlaProArgIleIleMetIleLeuTyHsHisLeuTyGlyAlaProIleGlnAsnArg 260  
 Db 732 TGGGCCCCCGCATCATCATGATTCCTTTACCACTCTATGSGGCGCCATCCAGAACCGC 791  
 Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
 Db 792 TGGCTGTGTACACATCATGTCCGACATTCGCCAATGCTAGCCCTTCTGAAACACAGCCATC 851  
 Qy 281 AsnPhePheLeuTyHsHisLeuSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300  
 Db 852 AACTTCTCTCTACTCTCTCATCAGCAGCGGTTCGCACCATGGCAGCGCCGACGCTC 911  
 Qy 301 LysAlaPhePheLysCysGlnProValGlnPheTyHsHisAsnPheSer 320  
 Db 912 AAGCTTTCTCAAGTCCAGACCACTGTACAGTTCTACACCAATCACTTTCC 971  
 Qy 321 IleThrSerSerProThrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyH 340  
 Db 972 ATAACAAGTAGCCCTCGATCTCGCCGCAAACTCACACTGCATCAAGATGCTGTGTAC 1031  
 Qy 341 GlnTyHsPheLysGlnTyHsProIleLysSerArgAsnAspSerLysSerSerTyHsGln 360  
 Db 1032 CAGTATGACAAAATGGAACCTTATAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 1091  
 Qy 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
 Db 1092 TTTGAGATGCCATTGGAGCTTGTGTCATCATCTGTG 1127

## RESULT 6

ACF05275

ID ACF05275 standard; cDNA; 2198 BP.

XX AC ACF05275;

XX DT 06-NOV-2003 (first entry)

XX DE Human G-protein coupled receptor HGPBMY34 cDNA.

XX KW HGPBMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;

XX KW human; neuroprotective; nootropic; tranquilizer; antimigraine;

XX KW neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;

XX KW cytotatic; cardiac; hypotensive; antitanginal; analgesic; anorectic;

XX KW anti-HIV; antiasthmatic; osteopathic; uropathic; antiulcer; antiallergic;

XX KW gene therapy; gene; ss.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

XX FT CDS 1247..2165

XX FT /\*tag= a

XX FT /product= "HGPBMY34"

XX PN WO2003050256-A2.

XX PD 19-JUN-2003.

XX PF 06-DEC-2002; 2002WO-US039290.

XX PR 06-DEC-2001; 2001US-0338371P.

XX PA (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX PI Feder JN, Gopal S, Mintier GA, Ramanathan CS;

XX DR WPI; 2003-577295/54.

XX DR P-PSDB; ABR62521.

XX PT New nucleic acid molecule encoding a human G-protein coupled receptor,

XX PT HGPBMY34, useful for diagnosing, preventing or treating diseases

XX PT involving the receptor, for example Parkinson's disease, dementia,

XX PT asthma, hypertension or cancer.

XX

PS Claim 1; Fig 1A-B; 112pp; English.

XX

The present sequence is that of cDNA encoding human HGPBMY34, a newly identified G-protein coupled receptor (GPCR) belonging to the group of 'Class A' GPCRs and showing homology to the Pfam model 7 transmembrane receptor, rhodopsin family. The cDNA was isolated by screening expressed sequence tag databases using GPCR sequences. HGPBMY34, also referred to as GPCR-P14 and/or GPCR-145, is highly expressed in brain (amygdala, caudate nucleus, corpus callosum, hippocampus, thalamus, substantia nigra), spinal cord and pituitary, indicating an association in neurological systems and conditions. It is also expressed in the bone marrow and testis. The invention provides HGPBMY34 polynucleotides, polypeptides and antibodies, expression vectors, host cells and antisense molecules, methods for screening for modulators of HGPBMY34 activity and/or function, and methods for diagnosing, treating, preventing and screening for disorders and diseases associated with abnormal HGPBMY34 activity, including: a disorder related to aberrant G-protein coupled signalling; a disorder related to aberrant cell cycle regulation; neurological disorders; anxiety; headache; migraine; schizophrenia; manic depression; delirium; dementia; severe mental retardation and dyskinesias, such as Huntington's disease or Gilles de la Tourette's syndrome; Parkinson's disease; brain disorders; spinal cord disorders; affective disorders; neoplastic disorders; cardiovascular disorders; myocardial infarction; an immunological disorder; immune-related disorders; endocrinal diseases; growth disorders; neuropathic pain; obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis; psychosis; metabolic disorders; pituitary disorders; urinary retention; ulcers; allergies; or benign prostatic hypertrophy (all claimed)

XX SQ Sequence 2198 BP; 456 A; 693 C; 581 G; 468 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 6,38e-187 Length: 2198  
 Score: 1967.00 Matches: 372  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-10-689-832-20 (1-372) x ACF05275 (1-2198)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
 Db 1047 ATGGACACACGACGCCACCTCGCAGCCACACAGCTCGCTGTCTTGGTGTCCCGCGC 1106  
 Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyHsSerLeuLeuLeuCysLeuGly 40  
 Db 1107 TCGGCTGGCGCTTGGGTTTCGTGCCCGTGTCTACTACAGCCTCTTGTGTGCTCGGT 1166  
 Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60  
 Db 1167 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGTGTGGCAAGACAGAG 1226  
 Qy 61 SerSerTyHsAsnTyHsLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80  
 Db 1227 TCCTCTACACATATCTTTGGCCTCGCTGTGTGGCAGCATCTTGTCTCTTTTCATA 1286  
 Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
 Db 1287 GTGTTTGTGACTTCTCTTGGGAAGATTTTCATCTTGAACATGACAGATGCTCAGGTCCC 1346  
 Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
 Db 1347 GACAGATCATAGAGTGTGGAAATTTCTATCATCCACACCTCCATATGGATTACTGTGA 1406  
 Qy 121 ProLeuThrIleAspArgTyHsIleAlaValCysHisProLeuLysTyHsThrValSer 140  
 Db 1407 CCGTTAACCATTTGACAGGTATATCGCTGTGTGCCACCCCGCTCAAGTACCACACGGTCTCA 1466  
 Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyHsIleThrCysPheLeuThrSer 160

```
Db 1467 TACCACGCCGACCCGGAAGTCATTGTAAAGTGTATACATCAGCTGCTCTCCAGCAGC 1526
Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 1527 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 1586
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 1587 CAGTCTCATCTGGATCCACTGCTTCCACGCTTACTGGTGGCCCTGCTCCATCTCTTTC 1646
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
Db 1647 ATCTGAACCAATCATTTGTGTACAAGCTCAGGAGGAAGCAATTTGCTCTCCGTGGC 1706
Qy 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 1707 TACTCCAGCGGGAGACCAACCGGCATCTTGTTCACCATTTACCTCCATCTTTGCCACAT 1766
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 1767 TGGGCCCCCGCATCATCATGATCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 1826
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 1827 TGGCTGGTACACATCATGTCGGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 1886
Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 1887 AACCTTCTCTCTACTGCTTCTTCATCAGCAAGCGGTTCGACCACTGCGCAGCGCTC 1946
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 1947 AAGCTTTCTTCAAGTGCAGAGAACCTCTGTAAGTTCTACACCAATCATTAATCTTTC 2006
Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 2007 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGATC 2066
Qy 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360
Db 2067 CAGTATGACAAAATGGAACACCTATAAAAGTCGTAAATGACAGCAAAAGCTCTACCAG 2126
Qy 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372
Db 2127 TTTGAGATGCCATTGGAGCTTGTGTCACTCACTCTG 2162
```

## RESULT 7

ACF05276  
ID ACF05276 standard; cDNA; 1110 BP.

XX ACF05276;

XX 06-NOV-2003 (first entry)

XX Human G-protein coupled receptor HGRPMY34 variant cDNA.

XX HGRPMY34; G-protein coupled receptor; receptor; GPCR-P14; GPCR-145;

XX human; neuroprotective; nootropic; tranquilizer; antimigraine;  
XX neuroleptic; antianemic; antidepressant; anticonvulsant; antiparkinsonian;  
XX cytotatic; cardiac; hypotensive; antianginal; analgesic; anorectic;  
XX anti-HIV; antiasthmatic; osteopathic; uropathic; antiulcer; antiallergic;  
XX gene therapy; gene; ss.

XX Homo sapiens.

XX Key Location/Qualifiers

XX CDS 1..1110

XX FT /\*tag= a  
XX FT /product= "HGRPMY34 variant"

XX WO2003050256-A2.

XX 19-JUN-2003.

XX

PF 06-DEC-2002; 2002WO-US039290.

XX 06-DEC-2001; 2001US-0338371P.

XX (BRIM ) BRISTOL-MYERS SQUIBB CO.

XX Feder JN, Gopal S, Mintier GA, Ramanathan CS;

XX WPI; 2003-577295/54.

DR P-PSDB; ABR62522.

XX New nucleic acid molecule encoding a human G-protein coupled receptor,  
HGRPMY34, useful for diagnosing, preventing or treating diseases  
involving the receptor, for example Parkinson's disease, dementia,  
asthma, hypertension or cancer.

XX Claim 1; Fig 2A-B; 112pp; English.

XX The present sequence is that of cDNA encoding a variant of human  
HGRPMY34 containing a deletion of 3 amino acids of the protein sequence  
but expected to share at least some of the expression patterns and  
function of HGRPMY34. This is a newly identified G-protein coupled  
receptor (GPCR) belonging to the group of 'Class A' GPCRs and showing  
homology to the Pfam model 7 transmembrane receptor, rhodopsin family.  
HGRPMY34 (or GPCR-P14, GPCR-145) is highly expressed in brain (amygdala,  
caudate nucleus, corpus callosum, hippocampus, thalamus, substantia  
nigra), spinal cord and pituitary, indicating an association in  
neurological systems and conditions. It is also expressed in the bone  
marrow and testis. The invention provides HGRPMY34 polynucleotides,  
polypeptides and antibodies, expression vectors, host cells and antisense  
molecules, methods for screening for modulators of HGRPMY34 activity  
and/or function, and methods for diagnosing, treating, preventing and  
screening for disorders and diseases associated with abnormal HGRPMY34  
activity, including: a disorder related to aberrant G-protein coupled  
signalling; a disorder related to aberrant cell cycle regulation;  
neurological disorders; anxiety; headache; migraine; schizophrenia; manic  
depression; delirium; dementia; severe mental retardation and  
dyskinesias, such as Huntington's disease or Gilles de la Tourette's  
syndrome; Parkinson's disease; brain disorders; spinal cord disorders;  
affective disorders; neoplastic disorders; cardiovascular disorders;  
acute heart failure; hypotension; hypertension; angina pectoris;  
myocardial infarction; an immunological disorder; immune-related  
disorders; endocrinal diseases; growth disorders; neuropathic pain;  
obesity; anorexia; HIV infections; cancers; bulimia; asthma; osteoporosis  
; psychosis; metabolic disorders; pituitary disorders; urinary retention;  
ulcers; allergies; or benign prostatic hypertrophy (all claimed)

SQ Sequence 1110 BP; 253 A; 351 C; 220 G; 286 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 1.43e-184 Length: 1110  
Score: 1939.50 Matches: 369  
Percent Similarity: 99.19% Conservative: 0  
Best Local Similarity: 99.19% Mismatches: 0  
Query Match: 98.60% Indels: 3  
DB: Gaps: 1

US-10-689-832-20 (1-372) x ACF05276 (1-1110)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerSerLeuSerTrpTrpSerProGly 20

Db 1 ATGGAGCACACGACGCCACCTCGCAGCAACAGCTCGCTGTCTGGTGGTCCCCCGGC 60

Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40

Db 61 TCGGCTCGGGCTTGGGTTTCGGCCCGGTGCTACTACAGCCTCTTGTGTGCTCGGT 120

Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60

Db 121 TTACCAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGTGTGCAAGACAGAAG 180

Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80

```
Db 181 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATA 240
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 241 GTGTTGTGACATCTCTGTTGGAGATTTCATCTTGAACTGACAGATGCTCAGGTCCCC 300
Qy 101 AspLysIleLeuGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120
Db 301 GACAAATCATAGAGTGTGGAATTCTCATCCATCCACACCTCCATATGATTACTGTA 360
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 361 CCGTTAACCAATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACCACACGGTCTCA 420
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 421 TACCCAGCCCGCACCCGGAAGTCAATGTGAAGTGTATACATCACTGCTTCTTGACACG 480
Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 481 ATCCCTTATTACTGTTGGTCCCAACATCTGACTGAACTACATCAGCACCTCTGTGCAT 540
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 541 CACGTCTCTCATCTGGATCCACTGCTTCCCGTCTACCTGGTGCCTGCTCCATCTTCTTC 600
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
Db 601 ATCTTGAACTCAATCATTTGTGTAAAGCTCAGGAGGAAGCAATTTTGTCTCCGTGGC 660
Qy 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 661 TACTCCAGCGGGAAGACACCGCACTTGTTCACCATTTACCTCCATCTTGGCACACTT 720
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 721 TGGGCCCCCGCATCATCATGATTTTACCACCTCTATGGGCGGCCCATCCAGAACCGC 780
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 781 TGCTGTGACATCATCATGATCCGACATTCGCACTGCACTGATGCTTCTGAAACAGCGATC 840
Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 841 AACTTCTCTCTACTGCTTTCATCAGCAAGCGTTCGACCATGCGCAGCGCCGACGCTC 900
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCAATCACTTTTCC 960
Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 961 ATAAAGTAGTCCCTTGGATCTCGCGGCAAACTCACATGCAATCAAGATGCTGGGTAC 1020
Qy 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360
Db 1021 CAGTATGACAAAATGGA-----AAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 1071
Qy 361 PheGluAspAlaIleGlyAlaCysValIleLeu 372
Db 1072 TTTGAAGATGCCATTTGGAGCTTGTGTATCATCTCTG 1107
RESULT 8
ADD18120
ID ADD18120 standard; DNA; 2189 BP.
XX
AC ADD18120;
XX
XX 15-JAN-2004 (first entry)
XX
DE Human G-protein coupled receptor (GPCR) DNA sequence Seq ID6.
XX
```

```
XX G protein coupled receptor; GPCR; signal transduction pathway; G protein;
KW Alzheimer's disease; Parkinson's disease; diabetes; dwarfism;
KW colour blindness; retinal pigmentosa; asthma; depression; schizophrenia;
KW sleeplessness; hypertension; anxiety; stress; renal failure;
KW cardiovascular disorder; neural disorder; oncology disorder;
KW immune disorder; neuroprotective; gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1047..2157
XX /*tag= a
XX /product= "Human G-protein coupled receptor"
XX /transl_except= (2085..2091, aa:KPIKS)
XX
XX WO2003016478-A2.
XX
XX 27-FEB-2003.
XX
XX 15-AUG-2002; 2002WO-US026017.
XX
XX 20-AUG-2001; 2001US-0313658P.
XX 12-SEP-2001; 2001US-0318675P.
XX 30-OCT-2001; 2001US-0340703P.
XX 26-NOV-2001; 2001US-0333417P.
XX 06-DEC-2001; 2001US-0338367P.
XX 06-FEB-2002; 2002US-0355596P.
XX
XX (BRIM ) BRISTOL-MYERS SQUIBB CO.
XX
XX Feder JN, Ramanathan CS, Gopal S, Mintier GA;
XX WPI: 2003-278558/27.
XX P-PSDB; ADD18023.
XX
XX New nucleic acid, useful for manufacturing a medicament for preventing,
XX treating or ameliorating a medical condition e.g., neural disorder.
XX
XX Example 2; SEQ ID NO 6; 251pp; English.
XX
XX This invention relates to novel G protein coupled receptors (GPCRs) and
XX their encoding nucleotide sequences. Many medically significant
XX biological processes are mediated by proteins participating in signal
XX transduction pathways involving G proteins. GPCRs are one of the largest
XX receptor superfamilies known. These receptors are biologically important
XX and malfunction of these receptors results in diseases such as
XX Alzheimer's, Parkinson's, diabetes, dwarfism, colour blindness, retinal
XX pigmentosa and asthma. They are also involved in depression,
XX schizophrenia, sleeplessness, hypertension, anxiety, stress, renal
XX failure and other cardiovascular, neural, oncology and immune disorders.
XX A modulator of the GPCRs of the invention may have neuroprotective
XX activity whilst the sequences of the invention may be useful for gene
XX therapy. The invention may also be useful for manufacturing a medicament
XX for preventing, treating or ameliorating a medical condition. The present
XX sequence is the DNA sequence encoding a human GPCR of the invention.
XX
XX Sequence 2189 BP; 451 A; 691 C; 581 G; 466 T; 0 U; 0 Other;
```

## Alignment Scores:

```
Pred. No.: 3.64e-184 Length: 2189
Score: 1939.50 Matches: 369
Percent Similarity: 99.19% Conservativeness: 0
Best Local Similarity: 99.19% Mismatches: 0
Query Match: 98.60% Indels: 3
DB: 10 Gaps: 1
```

US-10-689-832-20 (1-372) x ADD18120 (1-2189)

```
Qy 1 MetGluHisThrHisAlaHisLeuAlaAsnSerSerLeuSerTrpTrpSerProGly 20
Db 1047 ATGGAGCACACACACGCCACCTCGCAGCAACAGCTCGCTGTCTGGTGGTCCCGGC 1106
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
```

Db 1107 TCGGCTCGGGCTTGGGTTCTGTCGCCGTGCTACTACAGCCTCTTGCTGTGCTCGGT 1166  
QY 41 LeuProAlaAsnIleLeuThrValIleLeuSerGlnLeuValAlaArgGlnIys 60  
Db 1167 TTACGAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAAGACAGAAG 1226  
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
Db 1227 TCTCTCTACAATATCTCTTGGCATCGCTGCTGCCGACATCTTGCTCTCTTTTCATA 1286  
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db 1287 GGTGTTGTGGACTTCTCTGTGGAGATTTCATCTTGAACATGCAGATGCTTCAGTCCCC 1346  
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
Db 1347 GACAAGATCATAGAAGTGTGGAAATCTCATCTCCATCCACACCTCATATGGATTACTGTA 1406  
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db 1407 CGGTTAACCATTTGACAGGTATATCGTGTGCGCACCTCCGCTCAAGTACCACACGCTCTCA 1466  
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 1467 TACCAGCCCGCACCGGNAAGTCATTGTAAGTGTTTACATCACCCTGTTCTGTGACACG 1526  
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
Db 1527 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 1586  
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db 1587 CAGTGCCTCATCTGATCCACTGCTTCCACCGTCTACCTGGTGGCTGTCTCTCTCTTC 1646  
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 1647 ATCTTGAACTCAATCATTTGTGACAGCTCAGAGAGAGAGCATTTCGTCTCGCTGGC 1706  
QY 221 TyrSerThrGlyLeuThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 1707 TACTCCAGGGGAGACACCCCATCTTGTTCACCATTTACCTCATCTTTGCCACACTT 1766  
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
Db 1767 TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 1826  
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
Db 1827 TGGCTGGTACACATCATGTCCGACATTGCCAACATGCTAGCCCTTCTGACACAGCCATC 1886  
QY 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db 1887 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGTTCCGACCATGGCAGCGCCACGCTC 1946  
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db 1947 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 2006  
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
Db 2007 ATACACAGTAGCCCTTGGATCTCGCGGCAAACTCACACTGCATGCATGCATGCTGTGATC 2066  
QY 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360  
Db 2067 CAGTATGACAAAATGGA-----AAAAGTCGTAATGACAGCAAAAGCTCTTACCAG 2117  
QY 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
Db 2118 TTGAGATGCCATTGGAGCTGTGTGCATCATCTG 2153

RESULT 9  
ADR16433

ID ADR16433 standard; cDNA; 2117 BP.  
XX AC ADR16433;  
XX DT 04-NOV-2004 (first entry)  
XX DE Human KOR3-like (KOR3L) cDNA.  
XX KW Human; KOR3-like; KOR3L; gene; ss; G-protein coupled receptor; GPCR;  
KW loss of balance; sensorimotor integration; obesity; anorectic.  
XX OS Homo sapiens.  
XX FH Location/Qualifiers  
FT CDS 280..1341  
FT /\*tag= a  
FT /product= "Human KOR3L protein"  
XX PN US2004161799-A1.  
XX PD 19-AUG-2004.  
XX PF 13-FEB-2004; 2004US-00779104.  
XX PR 14-FEB-2003; 2003US-0447447P.  
XX PR 14-AUG-2003; 2003US-0495577P.  
XX PA (MURP/) MURPHY A J.  
XX PA (CROLL/) CROLL-KALISH S.  
XX PI Murphy AJ, Croll-Kalish S;  
XX DR WPI; 2004-593074/57.  
XX DR P-PSDB; ADR16434.  
XX PT Identifying an agent capable of modulating a KOR3-like protein (a class  
PT of G-protein coupled receptor protein), useful for treating  
PT obesity, comprises contacting agent with KOR3like protein and detecting  
PT binding.  
XX PS Disclosure; SEQ ID NO 1; 16pp; English.  
XX CC The invention relates to a method for identifying an agent capable of  
CC modulating activity of a KOR3-like (KOR3L) protein (a class of G-protein  
CC coupled receptor protein (GPCR)) or protein fragment, involving  
CC contacting a test agent with a KOR3L protein or protein fragment and  
CC determining the ability of the test agent to bind the KOR3L protein or  
CC protein fragment. The invention also relates to a method of treating  
CC KOR3L-mediated loss of balance or sensorimotor integration by  
CC administering an agent capable of activating KOR3L, a method of treating  
CC obesity involving administering an agent capable of inhibiting KOR3L, a  
CC pharmaceutical composition comprising an agent identified by a method of  
CC the invention and a carrier, and a transgenic animal comprising a  
CC modification of an endogenous KOR3L gene. The methods of the invention  
CC are useful for identifying an agent capable of modulating a KOR3L  
CC protein. The agent is useful for treating KOR3L mediated loss of balance  
CC or sensorimotor integration, and for treating obesity. This sequence  
XX represents cDNA encoding the human KOR3L protein of the invention.  
SQ Sequence 2117 BP; 490 A; 625 C; 497 G; 505 T; 0 U; 0 Other;  
Alignment Scores: 6.33e-181 Length: 2117  
Pred. No.: 1907.00 Matches: 372  
Score: 83.22% Conservative: 0  
Percent Similarity: 83.22% Mismatches: 0  
Best Local Similarity: 96.95% Indels: 75  
Query Match: 13 Gaps: 1  
DB: 1  
US-10-689-832-20 (1-372) x ADR16433 (1-2117)  
QY 1 MetGluHisThrHisAlaHisLeuAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
|||||

Db 280 ATGAGCAGCAGCAGCCACCTCGCAGCAGCAACAGCTCGCTGCTTGGTGGTCCCGCGC 339  
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
Db 340 TCGGCTCGCGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 399  
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60  
Db 400 TTACCAAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 459  
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
Db 460 TCTCTCTACACTATCTCTGGCACTGGCTGGCTGGCACTCTGGTCTCTCTTTTTCATA 519  
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db 520 GTGTTTGTGGACTTCTCTGTGGAGATTTCACTTGAACATGCAGATGCTCAGGTCCCC 579  
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120  
Db 580 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACCTCCATATGATGATTAAGTGA 639  
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db 640 CCGTTAACCATTTGACAGGTATATCGCTGTGCGCCCGCTCAAGTACCACACGGTCTCA 699  
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 700 TACCCAGCCCGCACCCGGAAGTCATTGTAAGTGTTCATACCTGCTCTCTGACCAGC 759  
Qy 161 IleProTyrTyrTrpProAsnIleThrGluAspTyrIleSerThrSerValHis 180  
Db 760 ATCCCCATATTACTGTGGTGGCCCAACATCTGAGCTGAAGACTACATCAGCACCTCTGTGCAT 819  
Qy 181 HisValLeuIleThrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db 820 CAGTCTCTATCTGGATCCACTGCTTACCGTCTACCTGTGGTCCCTGCTCCATCTCTTC 879  
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 880 ATCTTGAACATCAATCTGTGTACAGCTCAGAGGAGAGCAATTTGCTCTCGTGGC 939  
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 940 TACTCCAGCGGAAGACCAACCGCATCTTGTTCACCATTACCTCCATCTTTGCCACACTT 999  
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
Db 1000 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 1059  
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuAsnThrAlaIle 280  
Db 1060 TGCTGGTGCATCATCATGTCGACATTCGACATGTAGCCCTTCTGNACACGCCATC 1119  
Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db 1120 AACCTTCTCTCTACTGCTTCTATCAGCAAGCGGTTCCGCACTATGGCAGCGCCACGCTC 1179  
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db 1180 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCCTACCAATCATTAACATTTTCC 1239  
Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
Db 1240 ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACACTGCATCAGATGCTGTGTATC 1299  
Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350  
Db 1300 CAGTATGACAAAATGGAACCTATAAAGTATCCCCGTGATTCATAGGTGTGGCAAC 1359  
Qy 350 ----- 350  
Db 1360 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGGTGTCCCATCTATGGCTGAGCAGC 1419

Qy 350 ----- 350  
Db 1420 TCTCCTTAAGAGTGTCTAATCCGATTTCTGTCTCCCGCAGACTGGCAATTTCTCAGACTG 1479  
Qy 350 ----- 350  
Db 1480 GTAGATGAGAAGAGATGGAAAGAGAAGAGAGCATGAAGCTTGTCTTTTACTTATGCA 1539  
Qy 351 -----SerArgAsnAspSerLysSerTyrGlnPheGluAspAlaIleG 366  
Db 1540 TTTATTTCCACAGAGTGTGTAATGACAGCAAAAGCTCTACCAGTTTGAAGATGCCATTGG 1599  
Qy 366 ValaCysValIleIleLeu 372  
Db 1600 AGCTTGTGTCAATCATCTG 1618  
RESULT 10  
AAS15731  
ID AAS15731 standard; cDNA; 1343 BP.  
XX AC AAS15731;  
XX 14-FEB-2002 (first entry)  
DT 14-FEB-2002 (first entry)  
XX  
DE DNA encoding chemokine receptor family related protein, NOV11.  
KW NOV; cytostatic; psoriasis; neutropenic; neuroprotectant;  
KW cerebroprotective; hepatic; antiinflammatory; diabetes; anaemia;  
KW haemostatic; atherosclerosis; gene therapy; neurogenesis; motility;  
KW differentiation; proliferation; haematopoiesis; wound healing;  
KW angiogenesis; cystic fibrosis; congenital myotonia; acute pancreatitis;  
KW haemophilia; allergy; Pendred syndrome; skeletal dysplasia;  
KW ischaemic injury; neuroepithelial disorder; hepatitis; heart failure;  
KW chemokine receptor; chromosome 1; ss.  
XX Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT 5'UTR 1  
FT CDS /tag= a  
FT /tag= b  
FT /product= "Chemokine receptor related protein, NOV11"  
FT sig\_peptide 2..142 /tag= c  
FT mat\_peptide 143..1060 /tag= d  
FT /note= "Mature chemokine receptor related protein, NOV11"  
FT 3'UTR 1064..1343 /tag= e  
XX WO200170978-A2.  
XX 27-SEP-2001.  
XX 20-MAR-2001; 2001WO-US009093.  
XX 20-MAR-2000; 2000US-0190768P.  
XX 22-MAR-2000; 2000US-0190835P.  
XX 22-MAR-2000; 2000US-0190972P.  
XX 22-MAR-2000; 2000US-0191199P.  
XX 24-MAR-2000; 2000US-0191947P.  
XX 28-MAR-2000; 2000US-0192657P.  
XX 28-MAR-2000; 2000US-0192664P.  
XX 28-MAR-2000; 2000US-0192665P.  
XX 28-MAR-2000; 2000US-0192984P.  
XX 29-MAR-2000; 2000US-0192836P.  
XX 31-MAR-2000; 2000US-0193843P.  
XX (CURA-) CURAGEN CORP.  
XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;



PI Vernet CAM;  
 XX WPI; 2001-639127/73.  
 DR P-PSDB; AAU10068.  
 XX  
 PT Polypeptides and nucleic acids related to chloride channel, insulin-like  
 PT growth factor family of proteins, useful for diagnosing and treating  
 PT cancer, cystic fibrosis, acute pancreatitis and Alzheimer's disease.  
 XX  
 PS Claim 9; Page 44; 151pp; English.  
 XX  
 CC The invention describes isolated NOVX (NOVX1-11) polypeptides. NOVX  
 CC polypeptides are useful for treating pathology associated with NOVX  
 CC polypeptides, determining the presence of or predisposition to a disease  
 CC associated with altered levels of NOVX, identifying agents binding to  
 CC NOVX and treatment of disorders associated with altered expression of  
 CC members of chloride channel-associated proteins e.g. cystic fibrosis and  
 CC congenital myotonia. NOVX proteins are useful in treatment of disorders  
 CC including psoriasis, cancer, diabetes, metabolic disorders of pancreas,  
 CC e.g. acute pancreatitis, abnormal growth and accumulation of mast cells  
 CC in one or more organs (e.g. haemophilia, anaemia), Fendred syndrome,  
 CC skeletal dysplasias, disorders characterised by altered cell shape,  
 CC motility, and apoptosis, ischaemic injury, hepatitis, neuroepithelial  
 CC disorders, hepatic disorders (e.g. cryptogenic cirrhosis) and in the  
 CC treatment of disorders of vascular smooth muscle cell differentiation,  
 CC (e.g. heart failure, stroke). NOVX nucleic acids and polypeptides are  
 CC useful to screen for molecules which inhibit or enhance NOVX activity or  
 CC function and are useful as targets for the identifying small molecules,  
 CC that modulate or inhibit e.g. neurogenesis, proliferation, motility, cell  
 CC differentiation, haematopoiesis, wound healing and angiogenesis. NOV  
 CC sequences are also useful for: identifying a cell or tissue type in a  
 CC biological sample; amplifying DNA sequences from very small biological  
 CC samples e.g. hair or skin or body fluids and as primers and probes to  
 CC identify and/or clone NOVX homologues. NOVX proteins are useful  
 CC immunogens to generate antibodies to monitor protein levels and modulate  
 CC NOVX activity. Cells comprising the nucleic acids are useful for  
 CC producing transgenic animals, for studying the function and/or activity  
 CC of NOVX protein and identifying and/or evaluating modulators of NOVX  
 CC protein activity. This sequence encodes NOVX1 (located on chromosome 1)  
 CC related to the chemokine receptor family of proteins, one of 12 NOV  
 CC polypeptides described in the method of the invention  
 XX  
 SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 8.52e-181 Length: 1343  
 Score: 1903.00 Matches: 371  
 Percent Similarity: 83.00% Conservative: 0  
 Best Local Similarity: 83.00% Mismatches: 1  
 Query Match: 96.75% Indels: 75  
 DB: 4 Gaps: 1

US-10-689-832-20 (1-372) x AAS15731 (1-1343)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
 DB 2 ATGGAGCACAGCAGCCACCTCGCAGCAACAGCTCGCTGTGTGGTGTCCCCCGGC 61  
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
 DB 62 TCGGCTCGCGCTTGGGTTTCGTGCGCTGTCTACTACAGCCCTTCTGTCTCGT 121  
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60  
 DB 122 TTACAGCAAAATATTGACAGTGATCATCTCTCCAGCTGGTGGCAGAGACAGAA 181  
 QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
 DB 182 TCCTCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATA 241  
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
 DB 242 GTGTTGTGGACTTCCTGTGTGGAGATTTTCATCTTGAACATGCGAGATGCGCTCAGGTCCCC 301

QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
 DB 302 GACAAGATCATAGAAAGTCTGGAAATTCATCCATCCACACCTCATATGATGATCTGTA 361  
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 DB 362 CCGTTAACATTGACAGGTATATCATCTGTCTGCCACCGCTCAAGTACCACACCGTCTCA 421  
 QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 DB 422 TACCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTTACATCACCTGCTCTCTGACCAGC 481  
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpGluAspTyrIleSerThrSerValHis 180  
 DB 482 ATCCCTCTATTACTGTGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 541  
 QY 181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 DB 542 CACGTCTCATCTGGATCCACTGCTTCAACGTCTACTGTGTGCGCTGCTCATCTTCTTTC 601  
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
 DB 602 ATCTTGAACCTCAATCAATTGTGTACAGCTCAGGAGGAAGAGCAATTTTCGTCTCGTGGC 661  
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIle 240  
 DB 662 TACTCCACGGGGAAGACACCGCCATCTTGTTCACCAATTACCTCCATCTTTGGCACACTT 721  
 QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
 DB 722 TGGGCCCCCGCATCATCATGATTTCTTACCACCTCTATGGGGCGCCCATCCAGAACCGC 781  
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
 DB 782 TGGCTGTGTACATCATGTCTCGACATTCGCCAACATGCTAGCCCTCTTGAACACAGCCATC 841  
 QY 281 AsnPhePheLeuTyrCysPheIleSerIysArgPheArgThrMetAlaAlaAlaThrIle 300  
 DB 842 AACTTCTTCTCTACTCTCTCATCAGCAAGCGGTTCGCGACCATGGCAGCGCCGCGCACGCTC 901  
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
 DB 902 AAGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTTC 961  
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 DB 962 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGTGTAC 1021  
 QY 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350  
 DB 1022 CAGTATGACAAAAATGGAAAACTATAAAAGTATCCCGGTGATTCCTAGGTGTGGCAAC 1081  
 QY 350 ----- 350  
 DB 1082 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGGTGTCCCATCCTATGCTGAGCAGC 1141  
 QY 350 ----- 350  
 DB 1142 TCTCCTTAAGAGTGCTAATCCGATTTCTGTCTCCCGCAGAGCTGGGCAATTTCTCAGACTG 1201  
 QY 350 ----- 350  
 DB 1202 GTAGATGAGAAGAGATGGAGAGAGAAAGAGAGCATGAAGCTGTGTTTTTACTTTATGCA 1261  
 QY 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366  
 DB 1262 TTTATTTTCCACAGAGTGTGTAATGACAGCAAAAGCTCCTACAGTTTGAAGATGCCATTGG 1321  
 QY 366 YAlaCysValIleIleLeu 372  
 DB 1322 AGCTTGTGTATCATCATCTG 1340



## RESULT 11

ADJ87768  
ID ADJ87768 standard; DNA; 1343 BP.

XX AC ADJ87768;

XX DT 06-MAY-2004 (first entry)

XX DE G-coupled protein receptor-related protein coding sequence #57.

XX novel protein; G-coupled protein receptor-related protein;  
KW cardiomyopathy; atherosclerosis; cell signal processing-related disorder;  
KW metabolic pathway modulation-related disorder; diabetes; cancer; stroke;  
KW Huntington's disease; epilepsy; anxiety; pain; hypercholesterolemia;  
KW obesity; hypertension; Crohn's disease; systemic lupus erythematosus;  
KW viral infections; bacterial infection; parasitic infection;  
KW hyperthyroidism; hypothyroidism; Von Hippel-Lindau syndrome;  
KW Alzheimer's disease; tuberculous sclerosis; hypercalcaemia; cerebral palsy;  
KW gene; ds.

XX Unidentified.

OS WO2002102321-A2.

PN 27-DEC-2002.

XX 18-JUN-2002; 2002WO-US019522.

XX 18-JUN-2001; 2001US-029899AP.

PR 18-JUN-2001; 2001US-0299134P.

PR 04-OCT-2001; 2001US-00972446.

PR 06-JUN-2002; 2002US-00299134.

PR 07-JUN-2002; 2002US-00298994.

XX (CURA-) CURAGEN CORP.

XX Anderson DW, Guo X, Gusev VY, Herrmann JL, Li L, Mezes PS;

PI Pena CEA, Spaderna SK, Zhong M;

XX WPI; 2003-167441/16.

DR P-PSDB; ADJ87769.

XX New MOLX polypeptides and polynucleotides, useful in gene therapy,

PT particularly for treating or preventing e.g. cardiomyopathy,

PT atherosclerosis, diabetes, adenoma, brain tumor, breast cancer, prostate

PT cancer, stroke or pain.

XX Claim 8; SEQ ID NO 203; 378pp; English.

XX The invention comprises the amino acid and coding sequences of novel G-

CC coupled protein receptor-related (MOL) proteins. The DNA and protein

CC sequences of the invention are useful for treating or preventing a MOL-

CC associated disorder, such as: cardiomyopathy, atherosclerosis, disorders

CC associated with cell signal processing and metabolic pathway modulation,

CC or diabetes. The DNA and protein sequences are also useful for the

CC treatment of: cancer, stroke, Huntington's disease, epilepsy, anxiety,

CC pain, hypercholesterolemia, obesity, hypertension, Crohn's disease,

CC systemic lupus erythematosus, viral infections, bacterial infections,

CC parasitic infections, hyperthyroidism, hypothyroidism, Von Hippel-Lindau

CC syndrome, Alzheimer's disease, tuberculous sclerosis, hypercalcaemia, or

CC cerebral palsy. The present DNA sequence encodes a MOL protein of the

## Alignment Scores:

Pred. No.:	8,52e-181	Length:	1343
Score:	1903.00	Matches:	371
Percent Similarity:	83.00%	Conservative:	0
Best Local Similarity:	83.00%	Mismatches:	1
Query Match:	96.75%	Indels:	75
DB:	10	Gaps:	1

SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

US-10-689-832-20 (1-372) x ADJ87768 (1-1343)

QY 1 MetGluHisThrHisAlaHisLeuAlaHisSerSerLeuSerTrpTrpSerProGly 20

Db 2 ATGGAGCACACGACGCGCCACCTGGCAGCCACAGCTCGTGTCTTGGTGGTCCCGCGC 61

QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40

Db 62 TCGGCTCGCGCTTGGGTTTCGTGCGCGTGTACTACAGCCTCTTGTGTGCTCGGT 121

QY 41 LeuProAlaAsnLeuLeuThrValLeuLeuLeuSerGlnLeuValAlaAArgGlnLys 60

Db 122 TTACAGCAATATCTTACAGTATCATCTCTCCAGCTGGTGGCAGAGACAGAG 181

QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80

Db 182 TCCTCTCACTATCTCTTGGCAGCTCGCTGCGCAGCATCTTGGTCTCTCTTTTCATA 241

QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100

Db 242 GTGTTTGTGGACTTCTCTTGGAAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCGC 301

QY 101 AspLysIleLeuGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120

Db 302 GACAAGATCATAGAGTGTGGAATTCATCTCCACACCTCCATATGGAATTAATGTA 361

QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140

Db 362 CGTTAACCATTTGACAGGTATATCACTGTCTGCCACCGCTCAAGTACCAACAGGTCTCA 421

QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160

Db 422 TACCAGCCGCGACCCGGAAGTCAATTGTAAGTGTATTACATCACCTGCTCTCTGACCAGC 481

QY 161 IleProTyrTyrTrpTrpProAsnIleThrGluAspTyrIleSerThrSerValHis 180

Db 482 ATCCCTTATTACTGGTGGCCCAACATCTGGAGCTGAAGACTACATCAGCACCTCTGTGCAT 541

QY 181 HisValLeuIleThrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200

Db 542 CAGTCTCTCATCTGGATCCACTGCTTACCGTCTACCTGGTGGCTGCTCCATCTCTTCT 601

QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220

Db 602 ATCTTGAATCAATCATTTGTGTACAGCTCAGGAGGAGAGCAATTTTCGTCTCGTGGC 661

QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240

Db 662 TACTCCAGGGGAAGACCCAGCCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 721

QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260

Db 722 TGGGCCCCCGCATCATCATGATTTCTTACCACCTCTATGGGGCGCCCATCCAGAACCGC 781

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280

Db 782 TGGCTGTGTACATCATGTCTCCAGCATTCGCCAACATGCTAGCCCTTCTGAACACACCCATC 841

QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300

Db 842 AACTTCTTCTCTACTGTCTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCACGCTC 901

QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320

Db 902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCAATACTTTTCC 961

QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340

Db 962 ATAACAGTAGTACCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGTAC 1021

QY 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350

Db 1022 CAGTATGACAAAAATGGAAAACTTATAAAGTATCCCGCGTGATTCATAGGTGTGGCAAC 1081

```
QY 350 ----- 350
Db 1082 TACTGCCTCTGTCTAATCCATTCCAGATGGAAGGTGCCATCCTATGCTGAGCAGC 1141
QY 350 ----- 350
Db 1142 TCTCCTTAAGAGTGCTAATCCGATTCTCTGCTCCCGCAGACTGGCAATTCTCAGACTG 1201
QY 350 ----- 350
Db 1202 GTAGATGAGAAGAGATGGAAGAGAAGAAGAGAGCATGAAGCTTGTTTTTACTTATGCA 1261
QY 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366
Db 1262 TTTATTTCCAGAGTCTGTATGACAGCAAAAGCTCTACCAGTTTGAGATGCCAATGG 1321
QY 366 yAlaCysValIleIleLeu 372
Db 1322 AGCTTGTGTCATCATCCTG 1340

RESULT 12
ADI79324
ID ADI79324 standard; DNA; 1343 BP.
XX
AC ADI79324;
DT 22-APR-2004 (first entry)
DE NOV11 coding sequence, SEQ ID 21.
KW Cytostatic; Immunosuppressive; Antiallergic; Antimicrobial; Vasotropic;
KW Respiratory; Hepatotropic; Virucide; Gastrointestinal; Antidiabetic;
KW Ophthalmologic; Antipsoriatic; Neuroprotective; Nootropic;
KW Antiarteriosclerotic; Hypotensive; Cardiant; Cerebroprotective;
KW Gene Therapy; NOVX; human; cancer; myelogenous leukaemia;
KW congenital neonatal autoimmune thrombocytopaenia; immunological disorder;
KW allergy; infection; asthma; lung disease; reproductive disorder;
KW haemangioma; deafness; liver cirrhosis; hepatitis C; gastric disorder;
KW diabetic retinopathy; psoriasis; multiple sclerosis; atherosclerosis;
KW hypertension; stroke; heart failure; chromosome 1; NOV11;
KW chemokine receptor; gene; ds.
XX Homo sapiens.
XX
XX Key Location/Qualifiers
FH CDS 2..1063
FT /*tag= a
FT /product= "NOV11"
XX
FN WO2004009635-A2.
XX
XX 29-JAN-2004.
XX
XX 04-OCT-2001; 2001WO-US031292.
XX
XX 20-MAR-2001; 2001US-00813432.
XX
XX (CURA-) CURAGEN CORP.
XX
XX Taupier RJ, Majumder K, Spaderna SK, Smithson G, Mezes PS;
XX Vernet CAM;
XX
XX WPI; 2004-123380/12.
XX P-PSDB; ADI79325.
XX
XX Claim 8; Page 43-44; 158pp; English.
XX
CC The present invention relates to novel NOVX proteins and their coding
CC sequences (ADI79304-ADI79327). The sequences are useful for the
CC manufacture of a medicament for treating a syndrome associated with a
CC human disease associated with the protein, or for diagnosing and treating
CC disorders associated with the NOVX protein, such as cancer, myelogenous
```

```
CC leukaemia, congenital neonatal autoimmune thrombocytopaenia,
CC immunological disorders, allergy and infection, asthma, lung diseases,
CC reproductive disorders, male and female reproductive diseases,
CC haemangioma, deafness, liver cirrhosis, hepatitis C, gastric disorders,
CC diabetic retinopathy, psoriasis, multiple sclerosis, atherosclerosis,
CC hypertension, stroke and heart failure. NOV11 represents a new member of
CC the chemokine receptor family and the gene is located on chromosome 1.
XX
SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 8.52e-181 Length: 1343
Score: 1903.00 Matches: 371
Percent Similarity: 83.00% Conservative: 0
Best Local Similarity: 83.00% Mismatches: 1
Query Match: 96.75% Indels: 75
DB: 12 Gaps: 1
```

US-10-689-832-20 (1-372) x ADI79324 (1-1343)

```
QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
Db 2 ATGGAGCACACGACGACGCCACCTCGCAGCAACAGCTCGCTGTCTTGGTGTCCCCCGGC 61
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 62 TCGGCCTGGGCTTGGGTTTGGTGGCGGTGCTACTACAGCCTCTTGTGTGCTCGGT 121
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60
Db 122 TTACAGCAAAATATCTTGCAGATGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 181
QY 61 SerSerTyrAsnTyrIleLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 182 TCCTCCCTACAACTATCTTGGCAGCTCGCTGCCGACATCTTGGTCCCTTTTTCATA 241
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 242 GTGTTTGTGACTTCTCTTGGAGATTTTCATCTTGAACATGACAGATGCCTCAGTCCCC 301
QY 101 AspLysIleLeuGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 302 GACAAAGATCATAGAAGTGCTGGAATTCTCATCTCCACACCTCCATATGATTACTGTA 361
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 362 CCGTTAACTTGAAGTATATCAGCTGTCTGCCACCCGCTCAAGTACACACAGGTCTCA 421
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 422 TACCCAGCCGACCCCGGAAGTCAATTGTAAGTGTTTTACATCACCTGCTTCTGACCCAGC 481
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 482 ATCCCTTATTACTTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 541
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrIleValProCysSerIlePhePhe 200
Db 542 CACGTCCTCATCTCGGATCCCATCTCCGCTTACCTGGTGGCTGCTGCTCCTCTTCTTC 601
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 602 ATCTTGAATCAATCATATTGTGTACAGCTCAGGAGGAGAGCAATTTTGTCTCCGTGGC 661
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 662 TACTCCAGGGGAAGACCAACCGCCATCTTGTTCACCATTTACCTCATCTTTGGCACACTT 721
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGCAATCATCATGATTTTATACACCTCTATGGGGGCCCATCCAGAACCGC 781
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
```

Db 782 TGCGTGGTACACATCATGTCGCAATTCGCAACATCTAGCCCTTCTGAAACACAGCCATC 841  
Qy 281 AenPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db 842 AACTTCTCTCTACTGCTTCATCAGCAGCGGTCCGACCATGGCAGCCGCCAGCTC 901  
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db 902 AAGGCTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATATAAATTTTCC 961  
Qy 321 IleThrSerSerProTyrPheSerProAlaAsnSerHisCysIleIleValTyr 340  
Db 962 ATAACAAGTAGCCCTCGGATCTCGCCGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1021  
Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350  
Db 1022 CAGTATGACAAAATGGAACCTATAAAGTATCCCCGTGATTCATAGGTGTGGCAAC 1081  
Qy 350 ----- 350  
Db 1082 TACTGCTCTGTCTAATCAATTTCCAGATGGGAAGGTGCCATCTATGGCTGAGCAGC 1141  
Qy 350 ----- 350  
Db 1142 TCTCCTTAAGAGTGTAAATCCGATTTCTCTCCGCGAGACTGGGCAATTCTCAGACTG 1201  
Qy 350 ----- 350  
Db 1202 GTAGATGACAGAGATGGAAGAGAGAGAGAGATGAAGCTGTTTTTACTTATGCA 1261  
Qy 351 -----SerArgAsnAspSerLysSerTyrGlnPheGluAspAlaIleG1 366  
Db 1262 TTTATTTCACAGAGTCGTAATGACAGCAAAAGCTCTACCAAGTTTGAAGATGCCATTGG 1321  
Qy 366 yAlaCysValIleIleLeu 372  
Db 1322 AGCTTGTGTATCATCCTG 1340

## RESULT 13

ID ADO56003 standard; cDNA; 1343 BP.

XX AC ADO56003;

XX 15-JUL-2004 (first entry)

XX DNA encoding human NOV11.

XX human; gene; ss; cancer; obesity; diabetic nephropathy;  
KW acute pancreatitis; stroke; multiple sclerosis.

XX Homo sapiens.

XX US2004058862-A1.

XX 25-MAR-2004.

XX 18-SEP-2002; 2002US-00246583.

XX 18-SEP-2002; 2002US-00246583.

XX (MAJU/) MAJUMDER K.

XX Majumder K;

XX WPI; 2004-268835/25.

XX P-PSDB; ADO56004.

XX Novel NOVX polypeptides, useful for treating cancer, obesity, diabetic  
PT nephropathy, acute pancreatitis, strokes and multiple sclerosis.

XX Disclosure; Page 28; 87pp; English.

XX

CC The invention relates to novel isolated NOVX nucleic acids and encoded  
CC polypeptides. The nucleic acids, polypeptides and antibodies raised  
CC against the polypeptides are useful for preventing or treating diseases  
CC associated with aberrant NOVX expression or activity e.g., cancer,  
CC obesity, diabetic nephropathy, acute pancreatitis, stroke, multiple  
CC sclerosis. The present sequence represents a NOVX nucleic acid of the  
CC invention.

XX SQ Sequence 1343 BP; 315 A; 403 C; 273 G; 352 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.: 8-52e-181 Length: 1343  
Score: 1903.00 Matches: 371  
Percent Similarity: 83.00% Conservative: 0  
Best Local Similarity: 83.00% Mismatches: 1  
Query Match: 96.75% Indels: 75  
DB: 12 Gaps: 1

US-10-689-832-20 (1-372) x ADO56003 (1-1343)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGly 20  
Db 2 ATGGAGCAGACGACGCCACCTCCAGCCCAACAGCTGCTGCTGTGGTGGTCCCCGGC 61  
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
Db 62 TCGGCTCGCGCTTGGGTTTGTGCGCTGGTCTACTACAGCTCTTGTGTGCTCGGT 121  
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60  
Db 122 TTACCAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGACAGAG 181  
Qy 61 SerSerTyrAsnTyrLeuLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
Db 182 TCCTCTAACAATATCTCTTGGCACTCGCTGCGCAATCTTGGTCTCTTTTTCATA 241  
Qy 81 ValPheValAspPheLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db 242 GTGTTTGGACTTCTGTTGGAAGATTCATCTTGAACATGCAGATGCCTCAGGTCCC 301  
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrVal 120  
Db 302 GACAGATCATAGAAGTGTGGAATTCATCATCCACACCTCCATATGGATTATGTA 361  
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db 362 CCGTTAACCATTCAGCAGGTATATCACTGTCTGCCACCCGCTCAAGTACCACAGGTCTCA 421  
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 422 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTATACATCACCTGCTTCTGACCAGC 481  
Qy 161 IleProTyrTyrTrpTrpProAsnIleThrPheGluAspTyrIleSerThrSerValHis 180  
Db 482 ATCCCTATTACTGGTGGCCCAACATCTGAGACTGAAGACTACATCAGCACCTCTGTGCAT 541  
Qy 181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db 542 CACGTCTCATCTGGATCCACTGCTTACCTTACCTGGTGGCCCTGCTTCCATCTTCTTC 601  
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 602 ATCTTGAACATCAATCATTTGTACAACTCAGGAGGAGAGCAATTTTCTGCTCTCGTGGC 661  
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 662 TACTCCAGGGGAGAGACCACGCCCATCTTGTTCACCATTTACCTCCATCTTTGCCACATT 721  
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
Db 722 TGGGCCCCCGCATCATCATGATTTTACCACCTCTATGGGGGCCCCCATCCAGAACCCGC 781

```
QY 261 TrpLeuValHisIleMetSerAspIleAlaAenMetLeuAlaLeuLeuAAsnThrAlaIle 280
Db |||||
QY 782 TGGCTGGTACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCCATC 841
Db |||||
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300
Db |||||
QY 842 AACTTCTTCTTACTGCTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCACGCTC 901
Db |||||
QY 301 LysAlaPhePheLeuLysGlnProValGlnPheTyrThrAsnHisAAsnPheSer 320
Db |||||
QY 902 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 961
Db |||||
QY 321 IleThrSerSerProTrpIleSerProAlaAenSerHisCysIleLysMetLeuValTyr 340
Db |||||
QY 962 ATAACAAGTAGCCCTCGATCTCGCCGCAACTCACACTGCATCAAGATGCTGTGTGAC 1021
Db |||||
QY 341 GlnTyrAspLysAAsnGlyLysProIle-Lys----- 350
Db |||||
QY 1022 CAGTATGACAAATGGAACCTATAAAGTATATCCCGTGATTCATAGGTGGCAAC 1081
Db |||||
QY 350 ----- 350
Db |||||
QY 1082 TACTGCCTCTGTCTAATCCATTTCCAGATGGAGGTGTCCCATCTATGGCTGAGCAGC 1141
Db |||||
QY 350 ----- 350
Db |||||
QY 1142 TCTCCTTAAGAGTGTAAATCCGATTCCTGTCTCCCGCAGAGCTGGGCAATTCAGACTG 1201
Db |||||
QY 350 ----- 350
Db |||||
QY 1202 GTAGATGAGAGATGGAAGAGAGAAAGGAGACATGAAGCTTGTTTACTTATGCA 1261
Db |||||
QY 351 -----SerArgAAsnAspSerLysSerTyrGlnPheGluAspAlaIleG1 366
Db |||||
QY 1262 TTTATTTCCACAGATCGCTAATGACAGCAAAAGCTCTACCAAGTTTGAAGATGCCATTGG 1321
Db |||||
QY 366 yAlaCysValIleIleLeu 372
Db |||||
QY 1322 AGCTGTGTGCATCATCCTG 1340
Db |||||
RESULT 14
ADO28955
ID ADO28955 standard; cDNA; 2273 BP.
XX
AC ADO28955;
XX
DT 29-JUL-2004 (first entry)
XX
DE Human novel GPCR PGR3 polynucleotide, SEQ ID NO:54.
XX
KW G protein-coupled receptor; GPCR; drug screening; diagnosis;
KW transgenic mouse; neurological disorder; adrenal gland disorder;
KW colon disorder; intestinal disorder; cardiovascular disorder;
KW muscular disorder; blood disorder; immune disorder; bone disorder;
KW joint disorder; metabolic disorder; nutritive disorder; cancer;
KW kidney disorder; liver disorder; lung disorder; breast disorder;
KW ovary disorder; uterus disorder; prostate disorder; testis disorder;
KW skin disorder; stomach disorder; pancreas disorder; spleen disorder;
KW thymus disorder; thyroid disorder; antiparkinsonian; antiemetic;
KW cytostatic; antiinflammatory; vasotropic; antianginal; antiarrhythmic;
KW CNS; central nervous system; respiratory; antidiarrhoeic; antidiabetic;
KW virucide; hepatotropic; antibacterial; antianemic; antiseborrheic;
KW dermatological; antiulcer; antithyroid; anti allergic; anorectic;
KW immunosuppressive; nephrotropic; gene therapy; GPCR modulator; human;
KW gene; ss.
XX
OS Homo sapiens.
XX
XX WO2004040000-A2.
XX
PD 13-MAY-2004.
XX
XX 09-SEP-2003; 2003WO-US028226.
```

```
XX 09-SEP-2002; 2002US-0409303P.
PR 09-APR-2003; 2003US-0461329P.
XX (PRIM-) PRIMAL INC.
XX Gaitanaris GA, Bergmann JB, Gragerov A, Hohmann J, Li F;
PI Madisen L, McIlwain KL, Pavlova MN, Vassilatis D, Zeng H;
XX WPI; 2004-390329/36.
DR P-PSDB; ADO28954.
XX Novel mammalian G protein coupled receptors, useful for identifying
PT compounds that modulates diagnosing and treating disease condition
PT associated with GPCR dysfunction e.g. autoimmune diseases, angina
PT pectoris, Parkinson's disease.
XX Claim 13; SEQ ID NO 54; 542pp; English.
XX The invention relates to human and mouse G protein-coupled receptors
CC (GPCRs) and nucleic acids encoding them. The invention also relates to
CC sequences at least 90% identical to the GPCR proteins and nucleic acids
CC of the invention; methods of treating, preventing or diagnosing diseases
CC associated with GPCRs of the invention; methods of screening for
CC compounds useful in the treatment of GPCR-related diseases; a transgenic
CC mouse comprising a GPCR gene of the invention; a mouse comprising a
CC mutation in a GPCR transgene or in an endogenous GPCR gene; cells derived
CC from the transgenic mice; kits comprising several mice, each of which has
CC a mutation in a different GPCR gene of the invention; and kits comprising
CC probes which hybridise to GPCR polynucleotides of the invention. The
CC invention further discloses variants of the GPCR polypeptides and vectors
CC comprising a GPCR nucleic acid. The GPCR nucleic acids and proteins may
CC be used in the diagnosis, treatment or prevention of a wide variety of
CC diseases including neurological disorders (e.g., Alzheimer's disease,
CC depression, diabetic neuropathy, Parkinson's disease or schizophrenia);
CC disorders of the adrenal gland; disorders of the colon or intestine
CC (e.g., Crohn's disease, diarrhoea, food poisoning or irritable bowel
CC syndrome); cardiovascular disorders (e.g., angina, cardiac arrhythmia or
CC myocardial infarction); muscular disorders; blood disorders (e.g.,
CC anaemia or leukaemia); immune disorders (e.g., autoimmune disorders or
CC AIDS); bone and joint disorders (e.g., osteoarthritis, rheumatoid
CC arthritis, gout or osteoporosis); metabolic or nutritive disorders (e.g.,
CC obesity, enzyme deficiency-related diseases or vitamin deficiency-related
CC diseases); and disorders of the kidney, liver, lung, breast, ovary,
CC uterus, prostate, testis, skin, stomach, pancreas, spleen, thymus and
CC thyroid (e.g., cancers). The present sequence represents a GPCR-encoding
CC nucleic acid of the invention. Note: The full sequence data for this
CC patent did not form part of the printed specification; those sequences
CC not shown were obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 2273 BP; 503 A; 691 C; 549 G; 530 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,76e-180 Length: 2273
Scores: 1903.00 Matches: 371
Percent Similarity: 83.00% Conservative: 0
Best Local Similarity: 83.00% Mismatches: 1
Query Match: 96.75% Indels: 75
DB: 12 Gaps: 1
US-10-689-832-20 (1-372) x ADO28955 (1-2273)
QY 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerSerLeuSerTyrTrpSerProGly 20
Db 453 ATGGAGCACAGCACGCGCACCTCGCCAGCCACAGCTCGTGTCTGTGGTGGTCCCCCGGC 512
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 513 TCGGCTCGCGCTTGGGTTTCGTGCCCCGTGTCTACTACAGCCCTTCTGTGCTCGCTCGGT 572
QY 41 LeuProAlaAAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60
Db |||||
```

573 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTGTGGCAAGACAGAG 632  
Qy SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
Db TCCTCTTACAACTATCTCTTGCCACTCGCTGTGTCGCGACATCTTGTCCTCTTTTCATA 692  
Qy ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db GTGTTTGTGGACTCTCTGTGGAAGATTTTCATCTTGAACATGACAGATGCTCAGGTCCCG 752  
Qy AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal 120  
Db GACACAGATCATAGAGTGTGGAAATCTCATCTCATCATCATCATCTCTCTCTCTCTCTCT 812  
Qy ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db CCGTTAAACATTGACAGGTATATCACTGTCTGCCACCCGCTCAAGTACCACACGGTCTCA 872  
Qy TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db TACCCAGCCGCGACCCGGAAGTCATTGTAAGTGTTTTACATCACTCTCTCTCTCTCTCT 932  
Qy IleProTyrTyrTrpTrpProAsnIleTyrThrGluAspTyrIleSerThrSerValHis 180  
Db ATCCCTATTACTGGTGGCCCAACATCTGACCTGAAGACTACATCAGCACCTCTGTGCAT 992  
Qy HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db CAGTCTCTCATCTGGATCCACTGCTTCACGCTCTACTGTGGCTGCTCCATCTCTCTC 1052  
Qy IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220  
Db ATCTTGAACCTCAATCATCTGTGTACAAAGCTCAGGAGGAAGCAATTTGCTCTCCGTGC 1112  
Qy TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db TACTCCAGCGGGAAGACCCGCGCATCTGTTCACCATTACCTTCCATCTTTGCGACAT 1172  
Qy TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
Db TGGGCCCCCGCATCATCATGATCTTTTACCACCTCTATGGGGGCGCCATCCAGAACCG 1232  
Qy TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
Db TGGCTGTGTACATCATCTCGACATTCGCAACATGCTAGCCCTCTGAACACAGCCATC 1292  
Qy AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db AACTTCTCTCTACTGCTTCATCAGCAAGCGGTTCCGACCATGGCAGCCGCGCATC 1352  
Qy LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db AAGGCTTTCTTCAAGTCCGCAAGCACTGTACAGTTCTACCACTCATTAACCTTTTCC 1412  
Qy IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
Db ATAACAAGTAGCCCTCGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGGTGAC 1472  
Qy GlnTyrAspLysAsnGlyLysProIle-Lys----- 350  
Db CAGTATGACAAAATGGAAACCTATATAAAGTATCCCGTGATTCATAGGTGTGGCAAC 1532  
Qy ----- 350  
Db TACTGCCTCTGTCTAATCATTTCCAGATGGGAGGTGTCCCATCTATGGTGGACAGC 1592  
Qy ----- 350  
Db TCTCCTTAAGAGTGCTAATCCGATTTCTGTCTCCGCGAGACTGGGCAATTTCTCAGACTG 1652  
Qy ----- 350  
Db GTAGATGAGAGAGATGGAAGAGAAAGGAGAGCATGAAGCTTGTTTTACTTATGCA 1712

Qy 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG 366  
Db 1713 TTTATTTCACAGAGTCGTAATGACAGCAAAAGCTCTACCAGTTTGAAGATGCCATTGG 1772  
Qy 366 yAlaCysValIleIleLeu 372  
Db 1773 AGCTTGTGTCAATCATCTCTG 1791  
RESULT 15  
AAD46858  
ID AAD46858 standard; cDNA; 1526 BP.  
XX  
AC AAD46858;  
XX  
DT 27-JAN-2003 (first entry)  
XX  
DE Human 7TM domain receptor 65507 cDNA.  
XX  
KW Human; adenosine deaminase; seven transmembrane domain receptor; cancer;  
KW 7TM; glycoproteinase; immune disorder; IGA deficiency; allergy; arrhythmia;  
KW rheumatoid arthritis; diabetes; atherosclerosis; cardiovascular disorder;  
KW hypertension; ischaemic heart disease; obesity; myocardial infarction;  
KW endothelial cell disorder; Grave's disease; psoriasis; brain disorder;  
KW Parkinson's disease; Alzheimer's disease; haematopoietic disorder;  
KW cerebral oedema; metabolic disorder; liver disorder; platelet disorder;  
KW chromosome mapping; tissue typing; gene therapy; neuroprotective;  
KW cytostatic; anorectic; cardiant; haemostatic; gene; ss.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT CDS 139..1200  
FT /\*tag= a  
FT /product= "Human 65507 protein"  
FT /note= "This region is specifically claimed as SEQ ID NO:  
FT 12 in claim 1 of the specification"  
FT sig\_peptide 139..306  
FT /\*tag= b  
FT mat\_peptide 307..1197  
FT /\*tag= c  
FT /product= "Human mature 65507 protein"  
XX  
WO200274960-A2.  
XX  
PD 26-SEP-2002.  
XX  
PF 08-NOV-2001; 2001WO-US051427.  
XX  
PR 08-NOV-2000; 2000US-0246768P.  
PR 08-NOV-2000; 2000US-0246772P.  
PR 15-NOV-2000; 2000US-0249185P.  
XX  
PA (MILL-) MILLENNIUM PHARM INC.  
XX  
PI Leiby KR, Kapeller-Libermann R, Glucksman M;  
XX  
DR WPI: 2002-759898/82.  
DR P-PSDB; AAE29236.  
XX  
PT New 38650, 28472, 5495, 65507, 81588 or 14354 nucleic acid molecules,  
PT useful for diagnosing and treating cancer, immune, cardiovascular,  
PT hematopoietic, brain, pain, metabolic, liver or platelet disorders, and  
PT in pharmacogenomics.  
XX  
PS Claim 1; Fig 17; 178pp; English.  
XX  
CC The present invention relates to novel 38650, 28472, 5495, 65507, 81588  
CC or 14354 nucleic acid molecules which encode adenosine deaminase, glyco-  
CC protease or seven transmembrane domain (7TM) receptor family members.  
CC Sequences of the invention are useful in diagnosing and treating cancer  
CC or aberrant cellular proliferation and/or differentiation (e.g. colon or  
CC lung cancer), immune disorders (e.g. selective IgA deficiency, rheumatoid

CC arthritis, diabetes, allergy), heart or cardiovascular disorders (e.g.,  
 CC hypertension, atherosclerosis, arrhythmias, ischaemic heart disease,  
 CC myocardial infarction, thrombus) including endothelial cell disorders  
 CC (e.g. psoriasis, Grave's disease), haematopoietic disorders, brain  
 CC disorders (e.g. cerebral oedema, Parkinson's or Alzheimer's disease),  
 CC pain and metabolic disorders (e.g. obesity), liver disorders or platelet  
 CC disorders. They are also useful in screening assays, predictive medicine  
 CC (e.g. diagnostic assays, prognostic assays, monitoring clinical trials  
 CC and pharmacogenetics) and prophylactic and therapeutic methods. The  
 CC nucleic acids may also be used in chromosome mapping, tissue typing and  
 CC forensic biology and as surrogate markers. Sequences of the invention are  
 CC also used in gene therapy. The present sequence is human 7TM domain  
 CC receptor 6507 cDNA

XX Sequence 1526 BP; 343 A; 463 C; 345 G; 374 T; 0 U; 1 Other;  
 SQ

## Alignment Scores:

Pred. No.: 2, 6e-178 Length: 1526  
 Score: 1879.00 Matches: 366  
 Percent Similarity: 82.33% Conservative: 2  
 Best Local Similarity: 81.88% Mismatches: 4  
 Query Match: 95.53% Indels: 75  
 DB: 6 Gaps: 1

US-10-689-832-20 (1-372) x AAD46858 (1-1526)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
 DB 139 ATGGAGCACACGACGCCACCTCGACGACACAGCTCGTCTGTGGTGGTCCCGCGC 198  
 QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
 DB 199 TCGGCTCGGGCTTGGGTTCGTGCGGTGCTACTACAGCCTCTTGACGTGCTCGGT 258  
 QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAArgGlnLys 60  
 DB 259 TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 318  
 QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAlaAspIleLeuValLeuPhePheIle 80  
 DB 319 TCCTCTACACTATCTCTGGCACTCGTGTGCGGACATCTTGGTCTCTCTTTTCATA 378  
 QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuLeuAsnMetGlnMetProGlnValPro 100  
 DB 379 GTGTTGTGGACTTCTCTGTGAAGATTTCATCTTGAACATGCAGATGCCCTCAGGTCCC 438  
 QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
 DB 439 GACAAGATCATAGAAGTGTGGAATTCATCTCATCCATCCACCTCCATATGATGATTACTGA 498  
 QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
 DB 499 CCGTTAACCATTTGACAGGTATATCGTCTGTGCCACCGCTCAAGTACACACGGTCTCA 558  
 QY 141 TyrProAlaAArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
 DB 559 TACCAGCCCGCACCCGGAAGTCATTGTAGTGTATTACATCACCTGCTCTCTGACCCAGC 618  
 QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
 DB 619 ATCCCTTATTACTGTGGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 678  
 QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
 DB 679 CACGTCTCTCATCTGGGTCTCATCTTCCCGTCTACCGGTCTACCGGTGCCCTGTCTCTTCTC 738  
 QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgLys 220  
 DB 739 ATCTTGAACCTCAATCGTTGTGTACAGCTCAGGAGGAAGCAATTTTCGTCTGTGTGGC 798  
 QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
 DB 799 TACTCCAGGGGAAGACCACCGCCATCTGTTCACCAATTACCTCCTCTTTGCCATCTT 858

QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
 DB 859 TGGCCCCCGCATCATCATGATTTTACACCTCTATGGGGGCCCATCCAGAACCGC 918  
 QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
 DB 919 TGGTGGTACACATCATGTCCGACATTTGCCACATGCTAGCCCTTCTGAACACAGCCATC 978  
 QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrIleu 300  
 DB 979 AACTTCTTCTCTACTGTTTCATCAGCAAGCGGTTCGCGCACCATGGCAGCGCCACGCTC 1038  
 QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
 DB 1039 AAGCTTTCTTCAAGTCCGACGAACCTGTACAGTTCTACACCAATCATAACTTTTTC 1098  
 QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
 DB 1099 ATAACAAGTAGCCCTCGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGTATC 1158  
 QY 341 GlnTyrAspLysAsnGlyLysProIle-Lys- 350  
 DB 1159 CAGTATGACAAAAATGAAAAACCTATAAAGTATCCCGGTGATTCCCATAGGTGGCAAC 1218  
 QY 350 ----- 350  
 DB 1219 TACTGCCTGTCTAAATCCATTTCCAGATGGGAAGGTGTCCCATCTCTATGGCTGAGCAGC 1278  
 QY 350 ----- 350  
 DB 1279 TCTCCTTAAGAGTCTAATCCGATTTCTGTCTCCCGCAGACTGGGCAATTTCTCAGACTG 1338  
 QY 350 ----- 350  
 DB 1339 GTAGATGAGAAGAGATGGAAGAGAAGAAAGGAGAGCATGAAGCTTGTTTTACTTATGCA 1398  
 QY 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366  
 DB 1399 TTTATTTTCCACAGAGTCGTAAATGACAGCAAAAGCTCTACCAAGTTTGAAGATGCCATTGG 1458  
 QY 366 YAlaCysValIleIleLeu 372  
 DB 1459 AGCTTGTGTATCATCTCTG 1477

Search completed: February 12, 2005, 13:49:36  
 Job time : 581 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:19:58 ; Search time 41 Seconds  
(without alignments)  
872.991 Million cell updates/sec

Title: US-10-689-832-20  
Perfect score: 1967  
Sequence: 1 MEHTAHILANSSLSWSPG.....NDSKSSYOFDAIGACVIL 372

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79: \*  
1: pir1: \*  
2: pir2: \*  
3: pir3: \*  
4: pir4: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %		DB	ID	Description
		Match	Length			
1	242	12.3	411	2	I56444	thyrotrophin-relea
2	242	12.3	412	2	S23436	thyroliberin recep
3	241	12.3	393	2	A39251	thyrotrophin-releas
4	238.5	12.1	398	2	JN0708	thyrotrophin-releas
5	231.5	11.8	355	2	A45177	chemokine (C-C) re
6	226	11.5	352	2	J80296	thyrotrophin-releas
7	225	11.4	357	2	JC7319	probable allatosta
8	225	11.4	423	2	JC7677	allatostatin recep
9	219.5	11.2	388	2	JN0605	somatostatin recep
10	219	11.1	384	2	A47249	brain-specific som
11	217.5	11.1	340	2	T18704	hypothetical prote
12	217.5	11.1	394	2	JC7209	gallatin receptor -
13	217.5	11.1	658	2	JC8011	G protein-coupled
14	217	11.0	367	2	JC2421	opioid receptor ho
15	217	11.0	367	2	I49022	kappa opioid recep
16	217	11.0	367	2	I56520	G protein-coupled
17	214.5	10.9	595	2	JC8012	G protein-coupled
18	213.5	10.9	355	2	I49339	macrophage inflam
19	213	10.8	352	2	A45747	neuropeptide Y/pep
20	213	10.8	370	2	S43087	orphan opioid rece
21	213	10.8	384	2	JC4629	somatostatin recep
22	212	10.8	360	2	T23063	hypothetical prote
23	211.5	10.8	376	2	T24368	hypothetical prote
24	211.5	10.8	380	2	JC2338	kappa opioid recep
25	211.5	10.8	504	2	T29338	hypothetical prote
26	210.5	10.7	380	2	S36143	kappa opioid recep
27	209	10.6	424	2	JH0164	neurotensin recep
28	208.5	10.6	477	2	JC7913	capa receptor (CGI
29	207.5	10.5	363	2	I57940	somatostatin recep

RESULT 1  
I56444  
thyrotrophin-releasing hormone receptor - mouse  
C:Species: Mus sp. (mouse)  
C:Date: 26-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 11-Jan-2000  
C:Accession: I56444  
R:Sellar, R.E.; Taylor, P.L.; Lamb, R.P.; Zabavnik, J.; Anderson, L.; Eidne, K.A.  
J. Mol. Endocrinol. 10, 199-206, 1993  
A:Title: Functional expression and molecular characterization of the thyrotrophin-relea  
A:Reference number: I56444; MUID:93249585; PMID:8387312  
A:Accession: I56444  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-411 <RES>  
A:Cross-references: GB:S60053; NID:g300151; PIDN:AAB26491.1; PID:g300152  
C:Superfamily: adenosine receptor A1

Query Match 12.3%; Score 242; DB 2; Length 411;  
Best Local Similarity 22.6%; Pred. No. 1.4e-12;  
Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;  
Qy 19 PGSCGLGF-VPVVYVYLLC-LGLPANILVTIILSQLVARRQKSSYN-YLLALAAADIL 75  
Db 16 PQVAVALEYQVVTILLVVVIGLGVINIMVVLVVR--TKHRTATNCYLVSADILM 73  
Qy 76 VLPFIVFDLLEDFILNMQMPQVDKI-----IEVLEPSSIHSTIWTYPL 122  
Db 74 VL-----VAAGLPNITDSIYGSWVYGVGCLCITYLQYLGINASSCSITAF 119  
Qy 123 TIDRYIAVCHPLKYTVSYPARTKRVIVSVYITCFLTSIPYVW-----PNIWT-EDVIS 176  
Db 120 TIERYIAICHPIKAQPLCTFSRAKILIFWA---FTSIYCMWFFLLDLNISTYKDAIV 176  
Qy 177 TSVHHVLWIHCFTYVLVPCSIPIFILNSIYVYKL-----PNIWT-EDVIS 210  
Db 177 ISCGYKISRNYSPIYLMDFGVYVMPMLATVLYGFIARILFLNPIPSDPKNSKTWKN 236  
Qy 211 ---PRKSNFRLRG-----YSTGKTTALTFTITSIPATLWAP-RIMILYHYLGAPI 257  
Db 237 DSTHQNKMNLTNTTRCNFSTVSSRKQVTKMLAVVVLIFALLWMPYRFLVVVNSFLSPFF 296  
Qy 258 QNRVLVHMSDIANMLALLNTAINFFLYCFISKPRMTMAAATLKAFFKCKQKQVQFYTNH 317  
Db 297 QENWFL-----FCRCIYILNSAINPVYINLMSQFR-----AARFKLCNCKQKPTKAAANY 348  
Qy 318 NFSITSS 324  
Db 349 SVALNYS 355

RESULT 2  
S23436

thyroliberin receptor - rat  
N:Alternate names: thyrotropin-releasing hormone receptor  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 13-Jan-1995 #sequence revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S23436; I53279; A49168; PQ0326  
R:de la Pena, P.; Delgado, L.M.; del Camino, D.; Barrios, F.  
Biochem. J. 284, 891-899, 1992  
A:Title: Cloning and expression of the thyrotropin-releasing hormone receptor from GH(3)  
A:Reference number: S23436; MUID:92322017; PMID:1377915  
A:Accession: S23436  
A:Molecule type: mRNA  
A:Residues: 1-412 <P>  
A:Cross-references: UNIPROT:Q01717; EMBL:X64630; NID:g57394; PIDN:CAA45913.1; PID:g57395  
R:Kimura, N.; Arai, K.; Sahara, Y.; Suzuki, H.; Kimura, N.  
Endocrinology 134, 432-440, 1994  
A:Title: Estradiol transcriptionally and posttranscriptionally up-regulates thyrotropin-  
A:Reference number: I53279; MUID:94102223; PMID:8275956  
A:Accession: I53279  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-412 <R>  
A:Cross-references: GB:D17469; NID:g464199; PIDN:BAA04289.1; PID:g464200  
R:Zhao, D.; Yang, J.; Jones, K.E.; Gerald, C.; Suzuki, Y.; Hogan, P.G.; Chin, W.W.; Tash  
Endocrinology 130, 3529-3536, 1992  
A:Title: Molecular cloning of a complementary deoxyribonucleic acid encoding the thyrotro  
A:Reference number: A49168; MUID:92283212; PMID:1317787  
A:Accession: A49168  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-12, 'D', 14-290, 'K', 292-412 <ZHA>  
A:Experimental source: GH cells  
A:Note: sequence extracted from NCBI backbone (NCBIN:104788, NCBI:P:104795)  
R:Iyama, M.; Monden, T.; Satoh, T.; Iizuka, M.; Murakami, M.; Iriuchijima, T.; Mori, M.  
Biochem. Biophys. Res. Commun. 184, 367-372, 1992  
A:Title: Differential regulation of thyrotropin-releasing hormone receptor mRNA levels b  
A:Reference number: PQ0326; MUID:92231953; PMID:1373613  
A:Accession: PQ0326  
A:Molecule type: mRNA  
A:Residues: 30-58, 'P', 60-222, 'T', 224-261 <YAM>  
A:Experimental source: strain Wistar  
A:Note: the authors translated the codon ACA for residue 88 as Ala  
C:Superfamily: adenosine receptor A1  
C:Keywords: G protein-coupled receptor; transmembrane protein  
F:1-22/Domain: transmembrane #status predicted <TM1>  
F:32-54/Domain: transmembrane #status predicted <TM2>  
F:71-92/Domain: transmembrane #status predicted <TM3>  
F:116-140/Domain: transmembrane #status predicted <TM4>  
F:165-186/Domain: transmembrane #status predicted <TM5>

Query Match 12.3%; Score 242; DB 2; Length 412;  
Best Local Similarity 22.6%; Pred. No. 1.4e-12;  
Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;  
19 PGSAGLGF-VPVVVYSLLC-LGLPANILTVIILSQLVARROKSSYN-VLLALAAADIL 75  
16 PQVALEYQVTVILLVVIIICGLGIVGNIMVLVVMR--TKHMTPTNCYLVS LAVADLM 73  
76 VLFPIVDFLEDFILNMOPQVPDKI-----IEVLEFSSIHSTSIWITVPL 122  
74 VL-----VAAGLPNITDSIYGSWVYGVGCLCITYLQYLGINASSCSITAF 119  
123 TIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYVW-----PNIWT-EDVIS 176  
120 TIERYIAICHPIKAQFLCTFSRAKKIIFVWA---FTSIYCMVLFLLDLNISTYKDAIV 176  
177 TSVHVLWIHCFTVYLPVCSIFFLNSIIVVKL-----210  
177 ISCGYKISRNYSPYILMDFGVYVPMILATVLYGFIARILFLNPISDPKSKWTKWN 236  
211 ---RKSNFRLRG-----YSTGKTTAILFTITSIFATLWAP-RIIMLYHYLGAPI 257  
237 DSTHQNKNMNLNTNRCFNSVTSRQVTKMLAVVILFALLWMPYRTL VVVNSFLSPFF 296

Query Match 12.3%; Score 241; DB 2; Length 393;  
Best Local Similarity 22.1%; Pred. No. 1.6e-12;  
Matches 81; Conservative 75; Mismatches 123; Indels 88; Gaps 15;  
19 PGSAGLGF-VPVVVYSLLC-LGLPANILTVIILSQLVARROKSSYN-VLLALAAADIL 75  
16 PQAAVALEYQVTVILLVVIIICGLGIVGNIMVLVVMR--TKHMTPTNCYLVS LAVADLM 73  
76 VLFPIVDFLEDFILNMOPQVPDKI-----IEVLEFSSIHSTSIWITVPL 122  
74 VL-----VAAGLPNITDSIYGSWVYGVGCLCITYLQYLGINASSCSITAF 119  
123 TIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYVW-----PNIWT-EDVIS 176  
120 TIERYIAICHPIKAQFLCTFSRAKKIIFVWA---FTSIYCMVLFLLDLNISTYKNAV 176  
177 TSVHVLWIHCFTVYLPVCSIFFLNSIIVV-----208  
177 VSCGYKISRNYSPYILMDFGVYVPMILATVLYGFIARILFLNPISDPKSKWTKWN 236  
209 -KLRRKSNFRLRG-----YSTGKTTAILFTITSIFATLWAP-RIIMLYHYLGAPI 257  
237 DSTHQNKNMNLNTNRCFNSVTSRQVTKMLAVVILFALLWMPYRTL VVVNSFLSPFF 296  
258 QNRWLHMSDIANMLALTAINFFLYCFISKRFRTMAAATLKAFKCKQKQPVQVYTNH 317  
297 QENWFLL---FCRICIYLSAINPVYINLMSQKFR---AAPRKLNCCKQKQKTEKAANY 348  
318 NFSITSS 324  
349 SVALNYS 355

Query Match 12.3%; Score 242; DB 2; Length 412;  
Best Local Similarity 22.6%; Pred. No. 1.4e-12;  
Matches 83; Conservative 71; Mismatches 125; Indels 88; Gaps 14;  
19 PGSAGLGF-VPVVVYSLLC-LGLPANILTVIILSQLVARROKSSYN-VLLALAAADIL 75  
16 PQVALEYQVTVILLVVIIICGLGIVGNIMVLVVMR--TKHMTPTNCYLVS LAVADLM 73  
76 VLFPIVDFLEDFILNMOPQVPDKI-----IEVLEFSSIHSTSIWITVPL 122  
74 VL-----VAAGLPNITDSIYGSWVYGVGCLCITYLQYLGINASSCSITAF 119  
123 TIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYVW-----PNIWT-EDVIS 176  
120 TIERYIAICHPIKAQFLCTFSRAKKIIFVWA---FTSIYCMVLFLLDLNISTYKDAIV 176  
177 TSVHVLWIHCFTVYLPVCSIFFLNSIIVVKL-----210  
177 ISCGYKISRNYSPYILMDFGVYVPMILATVLYGFIARILFLNPISDPKSKWTKWN 236  
211 ---RKSNFRLRG-----YSTGKTTAILFTITSIFATLWAP-RIIMLYHYLGAPI 257  
237 DSTHQNKNMNLNTNRCFNSVTSRQVTKMLAVVILFALLWMPYRTL VVVNSFLSPFF 296

258 QNRWLHMSDIANMLALTAINFFLYCFISKRFRTMAAATLKAFKCKQKQPVQVYTNH 317  
297 QENWFLL---FCRICIYLSAINPVYINLMSQKFR---AAPRKLNCCKQKQKTEKAANY 348  
318 NFSITSS 324  
349 SVALNYS 355

RESULT 3  
A39251  
thyrotropin-releasing hormone receptor - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 22-Jan-1993 #sequence\_revision 22-Jan-1993 #text\_change 09-Jul-2004  
C:Accession: A39251  
R:Straub, R.E.; Frech, G.C.; Joho, R.H.; Gershengorn, M.C.  
Proc. Natl. Acad. Sci. U.S.A. 87, 9514-9518, 1990  
A:Title: Expression cloning of a cDNA encoding the mouse pituitary thyrotropin-releasing  
A:Reference number: A39251; MUID:91088548; PMID:2175902  
A:Accession: A39251  
A:Molecule type: mRNA  
A:Residues: 1-393 <STR>  
A:Cross-references: UNIPROT:P21761; GB:M59811; GB:M37490; NID:g202153; PIDN:AAA40480.1;  
C:Superfamily: adenosine receptor A1  
C:Keywords: G protein-coupled receptor; glycoprotein; membrane protein

Query Match 12.3%; Score 241; DB 2; Length 393;  
Best Local Similarity 22.1%; Pred. No. 1.6e-12;  
Matches 81; Conservative 75; Mismatches 123; Indels 88; Gaps 15;  
19 PGSAGLGF-VPVVVYSLLC-LGLPANILTVIILSQLVARROKSSYN-VLLALAAADIL 75  
16 PQAAVALEYQVTVILLVVIIICGLGIVGNIMVLVVMR--TKHMTPTNCYLVS LAVADLM 73  
76 VLFPIVDFLEDFILNMOPQVPDKI-----IEVLEFSSIHSTSIWITVPL 122  
74 VL-----VAAGLPNITDSIYGSWVYGVGCLCITYLQYLGINASSCSITAF 119  
123 TIDRYIAVCHPLKHTVSPARTKRVISVYITCFLTSIPYVW-----PNIWT-EDVIS 176  
120 TIERYIAICHPIKAQFLCTFSRAKKIIFVWA---FTSIYCMVLFLLDLNISTYKNAV 176  
177 TSVHVLWIHCFTVYLPVCSIFFLNSIIVV-----208  
177 VSCGYKISRNYSPYILMDFGVYVPMILATVLYGFIARILFLNPISDPKSKWTKWN 236  
209 -KLRRKSNFRLRG-----YSTGKTTAILFTITSIFATLWAP-RIIMLYHYLGAPI 257  
237 DSTHQNKNMNLNTNRCFNSVTSRQVTKMLAVVILFALLWMPYRTL VVVNSFLSPFF 296  
258 QNRWLHMSDIANMLALTAINFFLYCFISKRFRTMAAATLKAFKCKQKQPVQVYTNH 317  
297 QENWFLL---FCRICIYLSAINPVYINLMSQKFR---AAPRKLNCCKQKQKTEKAANY 348  
318 NFSITSS 324  
349 SVALNYS 355

RESULT 4  
JN0708  
thyrotropin-releasing hormone receptor - human  
C:Species: Homo sapiens (man)  
C:Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: S40682; JN0759; S50151; S50152; I38356; JN0708  
R:Matre, V.; Karlsen, H.E.; Wright, M.S.; Lundell, I.; Fjeldheim, A.K.; Gabrielsen, O.S.  
Biochem. Biophys. Res. Commun. 195, 179-185, 1993  
A:Title: Molecular cloning of a functional human thyrotropin-releasing hormone receptor.  
A:Reference number: S40682; MUID:93371401; PMID:8395824  
A:Accession: S40682  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-398 <MAT>







Db 185 AIAAIAVWVILLASVPVLSHGVEVTTY-SAAHTACVFILEADINRPGYNKPVFOI 243  
QY 189 ---FTYLVPCSIFFILNSIIVYKLRKSNFRNGYSTG-----KTTAILFTI 233  
Db 244 IPPATSYVPLALICGLYLWLVRL-----WRGAAPGHVSAESRRGKRVTRVVVV 296  
QY 234 TSIFATLWAP-RIIMLYLHYGAPIONRWLHMSDIANMLALNTAINFFLYCFISKRP 292  
Db 297 VAIFAVCFWPIQLILVLKSDVKEITNTSV--MIQIVSHVLAVMSCVNPILYAFLSDFH 354  
QY 293 RTMAAATLKAFK---C-----QKOPQVFTNHNFSITSPWISPAHSHCKMLVQYDGN 345  
Db 355 R-----KAFKRVINCAGQAQAPGRY-----HRASTIQOQPOAN 389  
QY 346 GKPIKS---RNDSSKS 357  
Db 390 GRALNNECVENDNKS 404

RESULT 9  
somatostatin receptor 4 - human  
JN0605  
C.Species: Homo sapiens (man)  
C.Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
C.Accession: JN0605; JN0762; A47457  
R.Xu, Y.; Song, J.; Bruno, J.P.; Berelowitz, M.  
Biochem. Biophys. Res. Commun. 193, 648-652, 1993  
A.Title: Molecular cloning and sequencing of a human somatostatin receptor, hSSTR4.  
A.Reference number: JN0605; MUID:93290656; PMID:8512564  
A.Accession: JN0605  
A.Molecule type: DNA  
A.Residues: 1-388 <X>  
R.Yamada, Y.; Kagimoto, S.; Kubota, A.; Yasuda, K.; Masuda, K.; Someya, Y.; Ihara, Y.; I  
Biochem. Biophys. Res. Commun. 195, 844-852, 1993  
A.Title: Cloning, functional expression and pharmacological characterization of a fourth  
A.Reference number: JN0762; MUID:93384611; PMID:8373420  
A.Accession: JN0762  
A.Molecule type: DNA  
A.Residues: 1-388 <YAM>  
A.Cross-references: GB:D16826; NID:G693907; PIDN:BA04106.1; PID:G693908  
R.Nohrer, L.; Raulf, F.; Bruns, C.; Buettner, R.; Hofstaedter, F.; Schule, R.  
Proc. Natl. Acad. Sci. U.S.A. 90, 4196-4200, 1993  
A.Title: Cloning and characterization of a fourth human somatostatin receptor.  
A.Reference number: A47457; MUID:93248256; PMID:8483934  
A.Accession: A47457  
A.Molecule type: DNA  
A.Residues: 1-82, 'T', 84-364, 'K', 366-388 <ROH>  
A.Cross-references: GB:L07833; NID:G307429; PIDN:AAA60565.1; PID:G307430  
A.Note: sequence extracted from NCBI backbone (NCBIN:130856, NCBI:130858)  
C.Comment: This protein mediates the diverse actions of the tetradecapptide somatostatin.  
C.Genetics:  
A.Gene: GDB:SSTR4  
A.Cross-references: GDB:202662; OMIM:182454  
A.Map position: 20p11.2-20p11.2  
A.Introns: #status absent  
C.Superfamily: vertebrate rhodopsin  
C.Keywords: G protein-coupled receptor; glycoprotein; hormone receptor; lipoprotein; pho  
F:47-73/Domain: transmembrane #status predicted <TM1>  
F:84-109/Domain: transmembrane #status predicted <TM2>  
F:121-142/Domain: transmembrane #status predicted <TM3>  
F:162-184/Domain: transmembrane #status predicted <TM4>  
F:208-238/Domain: transmembrane #status predicted <TM5>  
F:257-284/Domain: transmembrane #status predicted <TM6>  
F:291-314/Domain: transmembrane #status predicted <TM7>  
F:24/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:119-198/Disulfide bonds: #status predicted  
F:161,253/Binding site: phosphate (Ser) (covalent) (by cAMP-dependent kinases) #status pr  
F:327/Binding site: palmitate (Cys) (covalent) #status predicted

Query Match 11.2%; Score 219.5; DB 2; Length 388;  
Best Local Similarity 22.0%; Pred. No. 9.6e-11;  
Matches 72; Conservative 75; Mismatches 130; Indels 51; Gaps 11;

QY 19 PGSACGLGFVPV-VYYSLLCLGLPANILTVILSQLVARRQKSSNYLLAALADILVL 77  
Db 39 PGDARAAGWVAIQCIYALVCLVGLGNALVIFVILR-YAKMKTATNIYLLNLAVALDEL-- 95  
QY 78 PFIVFVDFLLEDFIL-NMQMPQVPDKIIEVLEFSSIHSTIWIPTLTIDRYTAVCHPLKY 136  
Db 96 -FMLSVPFVASSAALRHPFGSVLCRAVLSVDGLNMTSVFCLTVLSVDVRYVAVVHPLRA 154  
QY 137 HTVSYPARTRKRVIVSVITCPLTSTPY-----YMWPN-IWTEBDYIST 177  
Db 155 ATYRPSVAKLNLGVGWLASLLVLPFAIPADTRPARGQAVACNLQWPHPAWAVFV-- 212  
QY 178 SVHVLVIWHCFV-YLVPCSIFFILNSIIVYKLR--KSNFRLRGYSTGTTAILFTI 233  
Db 213 -----VYTFLLGFLPLVLAIGLYLLGVKRAVALRAGWQRRSEKKITRLVLVW 264  
QY 234 TSIFATLWAPRIIMLYLHYGAPIONRWLHMSDIANMLALNTAINFFLYCFISKRP 293  
Db 265 VVVFVLCWMPFYVQLNLVVTSLD-----ATVNHVSLTLYANSANPILYGLFLSDNFR 319  
QY 294 -----TWAAATLKAFKCKQKQPVQY 314  
Db 320 RSFQVLCIRCLLEGAGGABEPLDYY 347

RESULT 10  
A47249  
brain-specific somatostatin receptor SSTR-4 - rat  
C.Species: Rattus norvegicus (Norway rat)  
C.Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
C.Accession: A47249  
R.Bruno, J.F.; Xu, Y.; Song, J.; Berelowitz, M.  
Proc. Natl. Acad. Sci. U.S.A. 89, 11151-11155, 1992  
A.Title: Molecular cloning and functional expression of a brain-specific somatostatin r  
A.Reference number: A47249; MUID:93087484; PMID:1360663  
A.Accession: A47249  
A.Status: preliminary  
A.Molecule type: nucleic acid  
A.Residues: 1-384 <BRU>  
A.Cross-references: UNIPROT:P30937; GB:M9544; NID:G207072; PIDN:AAA42180.1; PID:G207073  
A.Note: sequence extracted from NCBI backbone (NCBIN:119731, NCBI:119732)  
C.Superfamily: vertebrate rhodopsin  
C.Keywords: G protein-coupled receptor; transmembrane protein

Query Match 11.1%; Score 219; DB 2; Length 384;  
Best Local Similarity 20.8%; Pred. No. 1e-10;  
Matches 75; Conservative 80; Mismatches 139; Indels 66; Gaps 11;

QY 3 HTHAHLAANSLSWMSPG-----SACGLGFVPV-----YYSLLCLGLPAN 44  
Db 2 NTPATPLGGEDTTWPGINASWAPDEEDAVRSDGTGTAGNVTIQCIIYALVCLVGLVN 61  
QY 45 ILTVILSQLVARRQKSSNYLLAALADILVFPFVDFLLEDFILNMQMPQVPDKII 104  
Db 62 ALVIFVILR-YAKMKTATNIYLLNLAVALDELFLMSVPFVASAAA--LRHWPFGVLCRAV 118  
QY 105 EVLEFSSIHSTIWIPTLTIDRYTAVCHPLKYHTVSYPARTRKRVIVSVITCPLTSTPY- 163  
Db 119 LSVDFGLNMTSVFCLTVLSVDVRYVAVVHPLRAATYRRPSVAKLNLGVWLASLLVLPFA 178  
QY 164 -----YMWPN-IWTEBDYISTSVHHLWIWHCFV-YLVPCSIFFILNS 204  
Db 179 VPADTRPARGQAVACNLQWPHPAWAVFV-----YTFLLGFLPLVLAIGLCYL 228  
QY 205 IIVYKLR--KSNFRLRGYSTGTTAILFTITTSIFATLWAPRIIMLYLHYGAPIONRW 261  
Db 229 LIGKRAVALRAGWQRRSEKKITRLVLMVTVTVFLCWMFPYVQVQLNLPVTSLD--- 285  
QY 262 LVHMSDIANMLALNTAINFFLYCFISKRP-----TWAAATLKAFKCKQKQPVQY 314  
Db 286 --ATVNHVSLTLYANSANPILYGLFLSDNFRSRFQVLCIRCLLETGTGGABEPLDYY 343





**THIS PAGE BLANK (43-18)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:11:42 ; Search time 179 Seconds  
(without alignments)  
1064.209 Million cell updates/sec

Title: US-10-689-832-20  
Perfect score: 1967  
Sequence: 1 MEHTHAHLAANSSLSWSPG.....NDKSSYQFIDAIGACVIL 372

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03: \*  
1: uniprot\_eprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1857	94.4	353	2	Q6DWJ6
2	1623	82.5	388	2	Q8TDU8
3	1317	67.0	265	2	Q9BYT4
4	1273	64.7	383	2	Q7T2L1
5	979	49.8	186	2	Q8GSP2
6	771	39.2	371	2	Q7T2L2
7	737	37.5	315	2	Q8GSL3
8	737	37.5	462	2	Q7Z601
9	737	37.5	1464	2	Q8NGB0
10	713	36.2	365	2	Q7TQ99
11	262	13.3	54	2	Q80UC8
12	260.5	13.2	394	2	Q17478
13	250.5	12.7	398	1	TRFR_BOVIN
14	249.5	12.7	398	1	TRFR_SHEEP
15	249.5	12.7	440	2	Q7PV87
16	248	12.6	549	2	Q9VZW5
17	242	12.3	412	1	TRFR_RAT
18	241	12.3	393	1	TRFR_MOUSE
19	240.5	12.2	413	2	Q65YS4
20	240	12.2	404	2	Q8JF27
21	239	12.2	462	2	Q8IRF3
22	238.5	12.1	398	1	TRFR_HUMAN
23	236	12.0	488	2	Q717R7
24	235.5	12.0	395	1	TRFR_CHICK
25	234	11.9	478	2	Q717T5
26	233.5	11.9	354	2	Q6YST0
27	233	11.8	397	2	Q9DDR1
28	233	11.8	540	2	Q9W025
29	232	11.8	435	2	Q8SWR3
30	231.5	11.8	355	1	CKR1_HUMAN
31	231.5	11.8	355	2	Q9MYJ8

32	231	11.7	397	2	Q8JFZ5	Q8jtz5 xenopus lae
33	230.5	11.7	344	2	Q7PHG9	Q7phg9 anopheles g
34	230.5	11.7	427	2	Q7RTK2	Q7rtk2 anopheles g
35	228	11.6	412	2	Q7LC54	Q7lc54 homo sapien
36	228	11.6	415	2	Q9GAM5	Q9gam5 homo sapien
37	228	11.6	415	2	Q9GZQ4	Q9gzq4 homo sapien
38	226	11.5	352	2	Q88820	Q88820 rattus norv
39	226	11.5	352	2	Q8R297	Q8r297 rattus norv
40	226	11.5	488	2	Q9W027	Q9w027 drosophila
41	225.5	11.5	382	2	Q9ERT2	Q9ert2 mus musculu
42	225	11.4	357	2	Q9NRC8	Q9nrc8 drosophila
43	225	11.4	423	2	Q964D4	Q964d4 periplaneta
44	223	11.3	346	2	Q6VYH4	Q6vyh4 heliothis z
45	223	11.3	352	2	Q9QWV3	Q9qwv3 rattus norv

ALIGNMENTS

RESULT 1  
Q6DWJ6 PRELIMINARY; PRT; 353 AA.  
AC Q6DWJ6;  
DT 25-OCT-2004 (Tremblrel. 28, Created)  
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
DE G protein-coupled receptor 139.  
GN Name=GPR139;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA Bonner T.I., Nagle J.W., Kauffman D.;  
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY635179; AAT65818.1; -  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor protein activity; IEA.  
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodpsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCRHDOPSND.  
DR PROSITE; PS50262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 353 AA; 40679 MW; CF1088FC85C84F85 CRC64;

Query Match	94.4%;	Score 1857;	DB 2;	Length 353;
Best Local Similarity	100.0%;	Pred. No. 1.5e-120;	Mismatches 0;	Indels 0; Gaps 0;
Matches 350;	Conservative 0;			
QY	1	MEHTHAHLAANSSLSWSPGSGACGLGFPVYVYISLLCLGLPANILTVIILSQLVARRQK	60	
Db	1	MEHTHAHLAANSSLSWSPGSGACGLGFPVYVYISLLCLGLPANILTVIILSQLVARRQK	60	
QY	61	SSNYLLAALAAADILVLFPIVDFLEDFILNMQMPQVDPKIEVLEPSSHTSITVTV	120	
Db	61	SSNYLLAALAAADILVLFPIVDFLEDFILNMQMPQVDPKIEVLEPSSHTSITVTV	120	
QY	121	PLTIDRYTAVCHPLKHYTVSPARTRKIVSVYITCFLTSIPYVWPNITWEDYISTSVH	180	
Db	121	PLTIDRYTAVCHPLKHYTVSPARTRKIVSVYITCFLTSIPYVWPNITWEDYISTSVH	180	
QY	181	HVLIWIHCFVTYVLPSCIFILNSIIYVKLRKSNFRLRGYSTGKTALTITTSIPATL	240	
Db	181	HVLIWIHCFVTYVLPSCIFILNSIIYVKLRKSNFRLRGYSTGKTALTITTSIPATL	240	
QY	241	WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALLNTAINFFLYCFISKRFRTHAAATL	300	
Db	241	WAPRIIMILYHLYGAPIQNRWLHVHMSDIANMLALLNTAINFFLYCFISKRFRTHAAATL	300	

QY 301 KAFFKQKQKQVQVQYTNHNFSTSPWISPA NSHCIRKMLVYQYDKNGKPIK 350  
|||||  
Db 301 KAFFKQKQKQVQVQYTNHNFSTSPWISPA NSHCIRKMLVYQYDKNGKPIK 350

## RESULT 2

Q8TDU8 PRELIMINARY; PRT; 388 AA.  
ID Q8TDU8  
AC Q8TDU8  
DT 01-JUN-2002 (TReMBLrel. 21, Created)  
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Putative G-protein coupled receptor.  
CN Name=GPCR;  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Takeda S., Kadowaki S., Haga T., Takasue H., Mitaku S.;  
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AB083594; BAB9307.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
SQ SEQUENCE 388 AA; 44893 MW; D8DF1CF0CC95D9D8 CRC64;

Query Match 82.5%; Score 1623; DB 2; Length 388;

Best Local Similarity 100.0%; Pred. No. 2.4e-104;  
Matches 308; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 43 ANILTVILSQLVARRQKSSNYLLALAAADILVLFVDFVDFLEDFILNQMPQVDPK 102  
|||||  
Db 78 ANILTVILSQLVARRQKSSNYLLALAAADILVLFVDFVDFLEDFILNQMPQVDPK 137  
|||||  
QY 103 IIEVLEFSSIHSTSWITVPLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCLTSIP 162  
|||||  
Db 138 IIEVLEFSSIHSTSWITVPLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCLTSIP 197  
|||||  
QY 163 YWPNINWTEYISTSVHVLIIWHCFVYLVPSCIFFILNSIIVYKLRRKSNFRLGYS 222  
|||||  
Db 198 YWPNINWTEYISTSVHVLIIWHCFVYLVPSCIFFILNSIIVYKLRRKSNFRLGYS 257  
|||||  
QY 223 TSKTTAILFTTISIFATLWAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALTAINF 282  
|||||  
Db 258 TSKTTAILFTTISIFATLWAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALTAINF 317  
|||||  
QY 283 FLYCFISKRFRTMAATLKAPFKCQKQVQVQYTNHNFSTSPWISPA NSHCIRKMLVYQY 342  
|||||  
Db 318 FLYCFISKRFRTMAATLKAPFKCQKQVQVQYTNHNFSTSPWISPA NSHCIRKMLVYQY 377  
|||||  
QY 343 DKNGKPIK 350  
|||||  
Db 378 DKNGKPIK 385  
|||||

## RESULT 3

Q9BYT4 PRELIMINARY; PRT; 265 AA.  
ID Q9BYT4  
AC Q9BYT4  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-JUN-2003 (TReMBLrel. 24, Last annotation update)  
DE Putative transmembrane receptor (Fragment).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=21853304; PubMed=11863363; DOI=10.1006/geno.2002.6711;  
RX Otolenghi C., Barbieri M., McElreavey K., Fellous M.;  
RT "Novel paralogy relations among human chromosomes support a link  
RT between the phylogeny of doublesex-related genes and the evolution of  
RT sex determination.";  
RL Genomics 79:333-343 (2002).  
DR EMBL; AJ303165; CAC33085.1; -.  
DR GO; GO:0016021; C:integral to membrane; IEA.  
DR GO; GO:0004872; F:receptor activity; IEA.  
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.  
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro; IPR000276; GPCR\_Rhodopsn.  
DR Pfam; PF00001; 7tm 1; 1.  
DR PRINTS; PR00237; GPCR\_Rhodopsn.  
DR PROSITE; PS0262; G\_PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor; Transmembrane.  
FT NON\_TER 1 265  
FT NON\_TER 265 265  
SQ SEQUENCE 265 AA; 30999 MW; 8F314368A492395B CRC64;  
Query Match 67.0%; Score 1317; DB 2; Length 265;  
Best Local Similarity 100.0%; Pred. No. 2.3e-83;  
Matches 252; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 43 ANILTVILSQLVARRQKSSNYLLALAAADILVLFVDFVDFLEDFILNQMPQVDPK 102  
|||||  
Db 14 ANILTVILSQLVARRQKSSNYLLALAAADILVLFVDFVDFLEDFILNQMPQVDPK 73  
|||||  
QY 103 IIEVLEFSSIHSTSWITVPLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCLTSIP 162  
|||||  
Db 74 IIEVLEFSSIHSTSWITVPLTIDRYIAVCHPLKHTVSPARTKRVIVSVYITCLTSIP 133  
|||||  
QY 163 YWPNINWTEYISTSVHVLIIWHCFVYLVPSCIFFILNSIIVYKLRRKSNFRLGYS 222  
|||||  
Db 134 YWPNINWTEYISTSVHVLIIWHCFVYLVPSCIFFILNSIIVYKLRRKSNFRLGYS 193  
|||||  
QY 223 TSKTTAILFTTISIFATLWAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALTAINF 282  
|||||  
Db 194 TSKTTAILFTTISIFATLWAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALTAINF 253  
|||||  
QY 283 FLYCFISKRFRT 294  
|||||  
Db 254 FLYCFISKRFRT 265  
|||||  
RESULT 4  
Q7T2L1 PRELIMINARY; PRT; 383 AA.  
ID Q7T2L1  
AC Q7T2L1  
DT 01-OCT-2003 (TReMBLrel. 25, Created)  
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)  
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)  
DE G protein-coupled receptor 142b (Fragment).  
GN Name=GPR142b;  
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;  
OC Tetraodontidae; Tetraodontidae; Takifugu.  
OX NCBI\_TaxID=31033;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22985413; PubMed=14623098; DOI=10.1016/S0014-5793(03)01196-7;  
RA Fredriksson R., Hoglund P.J., Gloriam D.B., Lagerstrom M.C.,  
RA Schloth H.B.;  
RT "Seven evolutionarily conserved human rhodopsin G protein-coupled  
RT receptors lacking close relatives.";  
RL FEBS Lett. 554:381-388(2003).  
RL EMBL; AY288414; AAP72123.1; -.



GO: GO:0016021; C: integral to membrane; IEA.  
GO: GO:0004872; F: receptor activity; IEA.  
GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR Pfam: PF00001; 7tm\_1; 1\_Rhodopsin.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAISHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLKALVHCKRQKQVQVYASHNFSTSPWISPAISHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor GPCR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).  
RL ENBL; AX255545; AK085057.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0004872; F: receptor activity; IEA.  
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAISHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLKALVHCKRQKQVQVYASHNFSTSPWISPAISHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor GPCR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).  
RL ENBL; AX255545; AK085057.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0004872; F: receptor activity; IEA.  
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKIVSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAISHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLKALVHCKRQKQVQVYASHNFSTSPWISPAISHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor GPCR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).  
RL ENBL;

GO: GO:0016021; C: integral to membrane; IEA.  
GO: GO:0004872; F: receptor activity; IEA.  
GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR Pfam: PF00001; 7tm\_1; 1\_Rhodopsin.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWHAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAHSHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLALVHCRKQKQVQVYASHNFSTSPWISPAHSHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor PGR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).  
RL ENBL; AX255545; AK085057.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0004872; F: receptor activity; IEA.  
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWHAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAHSHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLALVHCRKQKQVQVYASHNFSTSPWISPAHSHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor PGR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2003).  
RL ENBL; AX255545; AK085057.1; -.  
DR GO: GO:0016021; C: integral to membrane; IEA.  
DR GO: GO:0004872; F: receptor activity; IEA.  
DR GO: GO:0001584; F: rhodopsin-like receptor activity; IEA.  
DR GO: GO:0007186; P: G-protein coupled receptor protein signalin. . .; IEA.  
DR InterPro: IPR000276; GPCR\_Rhodopsin.  
DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00262; G-PROTEIN\_RECEP\_F1\_2; 1.  
KW Receptor.  
FT NON TER 1  
FT NON TER 383  
SQ SEQUENCE 383 AA; 42610 MW; 76180AA3B97EC6D1 CRC64;  
Query Match 64.7%; Score 1273.5; DB 2; Length 383;  
Best Local Similarity 66.8%; Pred. No. 3.2e-80;  
Matches 249; Conservative 34; Mismatches 66; Indels 25; Gaps 5;  
QY 1 MEHTAHLAANSLSWSPG-----SACGLGFVPPVYVYSLLCGLP-ANILTVILS 52  
DB 6 MEHSHLPSALLPNSSWSGQHPSEATQCPGLPFIYVYVYSLLCGLP-ANILTVILS 65  
QY 53 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 112  
DB 66 QLVARRQKSYNYLLAALAAADILVLFPIVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSI 125  
QY 113 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 172  
DB 126 HTSIWITVPLTDRIYAVCHPLKYHTVSPARTKVIUSVYITCFLTSPITPYWNIWTE 185  
QY 173 -----DYISTSVHVLWIHCFTVYLVPCSIFFILNSIIYVKL-RRKSNRL 218  
DB 186 LPQTGGGGGAGGNNRTVAQQLVWHAHCITVFLPCTVFFSLNAVIVHKRRRSCFRL 245  
QY 219 RGYSTGKTTAILFTITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 275  
DB 246 RGYSTGKTTAILLAITSIPATLWAPRIIMILYHLYGAPIQNWRW---LVHIMSDIANMLAL 305  
QY 276 LNTAINPFYCFISKFRFTWAAATLKAFKCKQKQVQVYTNHNFSTSPWISPAHSHCI 335  
DB 306 LNTGVNPFYCFISKFRFGMAANVLALVHCRKQKQVQVYASHNFSTSPWISPAHSHCI 365  
QY 336 KMLVQYQDKNGKPI 349  
DB 366 KMLVQYQDKNGKPI 379  
RESULT 5  
Q86SP2 PRELIMINARY; PRT; 186 AA.  
ID Q86SP2  
AC Q86SP2;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE G protein-coupled receptor PGR3 (Fragment).  
OS Homo sapiens (human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
ON NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;  
RA Vassiliadis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,  
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,  
RA Bergmann J.E., Galtanaris G.A.;  
RT "The G protein-coupled receptor repertoire of human and mouse.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908 (2

```
Db 92 IFVGFLETAVFHREVPALLRSVSAEFAANHASINSVAELTVDRYVALCHPLLHRQIS 151
Qy 141 YPARTRKIVSVYITCFILTSIPYWPNIWTDYISTSVHVLWIHCFYVLPSCSIFP 200
Db 152 YPARTRRIAMVLVLSASGLPFFWSDMRNSHPPTAILDAILWTHVTIYFLPCSIFL 211
Qy 201 ILNSIIIVKLARKSNFRL---RGYST-----GKTTAILFTITSIFATLWAPRIIMLY 250
Db 212 VLNSLIITHLRARQOQCSQDERGQSQAPPRLGKTTAMLAITSVFSVLWAPRTVVYI 271
Qy 251 HLYGAPIQNRWLHVHMSDIANMLALNTAINFFLYCFISKFRP 293
Db 272 HLYVSSVHRDWRVHLAYDLSNMLAMLTAVNFFLYCFVSKPFR 314

RESULT 7
Q86SL3 PRELIMINARY; PRT; 315 AA.
AC Q86SL3;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DE G protein-coupled receptor PGR2 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;
RX Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA Bergmann J.E., Gaftanaris G.A.;
RT "The G protein-coupled receptor repertoire of human and mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
DR EMBL; AY255622; AAO85134.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PRINTS; PR00237; GPCR_RHODOPSIN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
FT NON TER 1 1
FT NON TER 315 315
SQ SEQUENCE 315 AA; 35173 MW; C82593E331EB9023 CRC64;

Query Match 37.5%; Score 737; DB 2; Length 315;
Best Local Similarity 46.3%; Pred. No. 3.2e-43;
Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2;

Qy 21 SACGLGFVVPVYVYLLCLGLPANILTVIILSOLVARQKSNYVLLAALDILVLFPI 80
Db 30 SPCVAGVIPVYVYVLLGLGLPVSLTAVALARLATRRFSYVYLLTASDIIQVVI 89
Qy 81 VFVDFLEDFILNMQMPQVPDKIIEVLEFSSIHSTSWITVPLTIDRYIACHPLKHYTVS 140
Db 90 VFAGELQGVALARQVQAVVVRNTANILEFAANHASVIAILLTVDRYTALCHPLHRAAS 149
Qy 141 YPARTRKIVSVYITCFILTSIPYWPNIWTDYISTSVHVLWIHCFYVLPSCSIFP 200
Db 150 SPGRTRRAIAAVLSAALLTGIPFYWLDMDWRDTPRTLDELVKWAHCLTVYFIPCGVFL 209
Qy 201 ILNSIIIVKLARKSNFRLRGYSTGKTTAILFTITSIFATLWAPRIIMLYHYLGAPIQNR 260
Db 210 VTNSAIHRLRRRGRSGLOP-RVGKSTAILLGITTLFTLLWAPRVFVWMLYHYVAPVRD 268
Qy 261 WLVIHMSDIANMLALNTAINFFLYCFISKFR-TMAAATLKAFKFC 306
Db 269 WRVHLALDVANNVAMLHTAANFGLYCFVSKTFRATVRQVIHDAVLP 315
```

```
RESULT 8
Q72601 PRELIMINARY; PRT; 462 AA.
AC Q72601;
DT 01-OCT-2003 (TReMBLrel. 25, Created)
DT 01-OCT-2003 (TReMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE G protein-coupled receptor 142.
GN Name=GPR142;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22985413; PubMed=14623098; DOI=10.1016/S0014-5793(03)01196-7;
RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
RA Schioth H.B.;
RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
RT receptors lacking close relatives.";
RL FEBS Lett. 554:381-388(2003).
DR EMBL; AY288421; AAP72130.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; F:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopsn.
DR Pfam; PF00001; 7tm 1; 1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor.
SQ SEQUENCE 462 AA; 51106 MW; E9120B52372C75CA CRC64;

Query Match 37.5%; Score 737; DB 2; Length 462;
Best Local Similarity 46.3%; Pred. No. 4.4e-43;
Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2;

Qy 21 SACGLGFVVPVYVYLLCLGLPANILTVIILSOLVARQKSNYVLLAALDILVLFPI 80
Db 151 SPCVAGVIPVYVYVLLGLGLPVSLTAVALARLATRRFSYVYLLTASDIIQVVI 210
Qy 81 VFVDFLEDFILNMQMPQVPDKIIEVLEFSSIHSTSWITVPLTIDRYIACHPLKHYTVS 140
Db 211 VFAGELQGVALARQVQAVVVRNTANILEFAANHASVIAILLTVDRYTALCHPLHRAAS 270
Qy 141 YPARTRKIVSVYITCFILTSIPYWPNIWTDYISTSVHVLWIHCFYVLPSCSIFP 200
Db 271 SPGRTRRAIAAVLSAALLTGIPFYWLDMDWRDTPRTLDELVKWAHCLTVYFIPCGVFL 330
Qy 201 ILNSIIIVKLARKSNFRLRGYSTGKTTAILFTITSIFATLWAPRIIMLYHYLGAPIQNR 260
Db 331 VTNSAIHRLRRRGRSGLOP-RVGKSTAILLGITTLFTLLWAPRVFVWMLYHYVAPVRD 389
Qy 261 WLVIHMSDIANMLALNTAINFFLYCFISKFR-TMAAATLKAFKFC 306
Db 390 WRVHLALDVANNVAMLHTAANFGLYCFVSKTFRATVRQVIHDAVLP 436

RESULT 9
Q8NGBO PRELIMINARY; PRT; 1464 AA.
AC Q8NGBO;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Seven transmembrane helix receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RA Suwa M., Sato T., Okouchi I., Arita M., Futami K., Matsumoto S.,
RL Teutsumi S., Aburatani H., Asai K., Akiyama Y.,
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB065914; BAC06129.1; -.
DR Genew; HGNC:20088; GPR142.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopn.
DR Pfam; PF00001; 7tm.1.1.
DR PRINTS; PRO0237; GPCRHOPOSN.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
KW Receptor; Transmembrane.
SQ SEQUENCE 1464 AA; 156507 MW; PF8C63F8353ACF7E CRC64;

Query Match 37.5%; Score 737; DB 2; Length 1464;
Best Local Similarity 46.3%; Pred. No. 1.2e-42;
Matches 133; Conservative 57; Mismatches 95; Indels 2; Gaps 2;

QY 21 SAGLGFPVYVYSLLCGLPANILTVIILSQLVARROKSSVYLLALAAADILVLPFI 80
Db 1113 SPCVAGIIPVIYVSVLLSGLP-----VALARLAARTKPSYHYLLALTASDIVTQVII 1172

QY 81 VFVDFLEDFILNMOPQVDPKIIIEVLEPSSIHSTIWIPTIDRYIAVCHPLKHYTVS 140
Db 1173 VFAGFLQGLAVLARQVQAVVRTANILEFAANHASVWIAITLVDEYALCHPLHRAAS 1232

QY 141 YPARTRKIVSVITCFLTSIPYYWPNITWEDYSTSVHHVLIWHTCFTVYLPCSIFF 200
Db 1233 SPGRTRRAIAAVLSAALLTGIPFYWLDWVRDTPSRTLDVLRKWAHCLTVYFIPCGVFL 1292

QY 201 ILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATLWAPRIIMILYHLYGAPIQNR 260
Db 1293 VTNSAILHLRRGRSGGLQPVGVSTAILLGLTTLFLLWAPRVFVMLYHYMYVAPVHRD 1351

QY 261 MVVHMSDIANMLALINTAINFLYCFISKRRP-TMAAATLKAFKFC 306
Db 1352 WRVHLALDVANNVAMLHTAANFGLCYFSKTRATVRQVIHDAYLPC 1398

RESULT 10
QY 07TON9 PRELIMINARY; PRT; 365 AA.
AC Q7TON9;
DT 01-OCT-2003 (TremBLrel. 25, Created)
DT 01-OCT-2003 (TremBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
DE G protein-coupled receptor 142.
GN Name=Gpr142;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22985413; PubMed=14623098; DOI=10.1016/S0014-5793(03)01196-7;
RA Fredriksson R., Hoglund P.J., Gloriam D.E., Lagerstrom M.C.,
RA Schioth H.B.;
RT "Seven evolutionarily conserved human rhodopsin G protein-coupled
RL receptors lacking close relatives.";
RL FEBS Lett. 554:381-388(2003).
DR EMBL; AY288428; AAP72137.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0004872; F:receptor activity; IEA.
DR GO; GO:0001584; F:rhodopsin-like receptor activity; IEA.
DR GO; GO:0007186; P:G-protein coupled receptor protein signalin. . .; IEA.
DR InterPro; IPR000276; GPCR_Rhodopn.
DR Pfam; PF00001; 7tm.1.1.
DR PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; UNKNOWN_1.
DR PROSITE; PS0262; G_PROTEIN_RECEP_F1_2; 1.
```

```
KW Receptor.
SQ SEQUENCE 365 AA; 40758 MW; ACA8649245CDAE47 CRC64;

Query Match 36.2%; Score 713; DB 2; Length 365;
Best Local Similarity 46.8%; Pred. No. 1.6e-41;
Matches 138; Conservative 51; Mismatches 96; Indels 10; Gaps 4;

QY 21 SAGLGFPVYVYSLLCGLPANILTVIILSQLVARROKSSVYLLALAAADILVLPFI 80
Db 64 SPCVAGIIPVIYVSVLLSGLP-----VALARLAARTKPSYHYLLALTASDIVTQVII 117

QY 81 VFVDFLEDFILNMOPQVDPKIIIEVLEPSSIHSTIWIPTIDRYIAVCHPLKHYTVS 140
Db 118 VFVDFLQGLAVLARQVQAVVRTANILEFAANHASVWIAVFTVDYRNALCRPLHRAAS 177

QY 141 YPARTRKIVSVITCFLTSIPYYWPNITWEDYSTSVHHVLIWHTCFTVYLPCSIFF 200
Db 178 SPGRTRRAIAAVIGVTLTGIPFYWLDWVRDTPSRTMDKLRKWAHCLTVIFIPCNVFL 237

QY 201 ILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATLWAPRIIMILYHLYGAPIQNR 260
Db 238 VTNSAILHLRRKRGQGLRPL-VSKSTAILLGLTSLFALLWAPRIIVMLYHYLYVAPVHRD 296

QY 261 MVVHMSDIANMLALINTAINFLYCFISKRRP-TMAAATLKAFKFC--OKQPVQ 312
Db 297 WRVHLALDIANMLANLNTENVNFGLYCFISKTRATVRQVICDVMACALKSQPKQ 351

RESULT 11
QY 080UC8 PRELIMINARY; PRT; 54 AA.
AC Q80UC8;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE G protein-coupled receptor PGR3 (Fragment).
GN Name=Gpr139;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22584407; PubMed=12679517; DOI=10.1073/pnas.0230374100;
RA Vassilatis D.K., Hohmann J.G., Zeng H., Li F., Ranchalis J.E.,
RA Mortrud M.T., Brown A., Rodriguez S.S., Weller J.R., Wright A.C.,
RA Bergmann J.E., Gaitanaris G.A.;
RT "The G protein-coupled receptor repertoires of human and mouse.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:4903-4908(2003).
DR EMBL; AY255548; AAO85060.1; -.
DR MGD; MGI:2685341; Gpr139.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON_TER 1 1
FT NON_TER 54 54
SQ SEQUENCE 54 AA; 6082 MW; 5B7B44A140BF8FDD CRC64;

Query Match 13.3%; Score 262; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 4.4e-11;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 226 TTAIFTTTSIFATLWAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTA 279
Db 1 TTAIFTTTSIFATLWAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTA 54

RESULT 12
QY 017478 PRELIMINARY; PRT; 394 AA.
AC Q17478;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TremBLrel. 26, Last annotation update)
```



QY 133 PLKHYTVSPARTKRVISVYITCFLTSIPYYW-----PNIWT-EDYISTSVHVLWI 186  
 DB 130 PIKAQPLCTFSRAKKIIFVMA--FTSIYCLMFLDLNISTYKDALVWSGGYKISR 186  
 QY 187 HCFTVYLVPSCFFILNSIIVKL-----RRKSN----- 215  
 DB 187 YSPYILMDFGVYVPMILATVYGFIAIILFLNIPSDPKNSNMWKNDSHTQKNLN 246  
 QY 216 -----FRLRGYSTGKTTAILFTITSIPATLWAP--RIIMLYHYLYGAPIQNRWLHIMS 267  
 DB 247 SKTSNRYFNSTVSSRKQVTKLAVVILFALLWMPVTLVVNSFLSPFQENWELL--- 303  
 QY 268 DIANMLALLNTAINFLYCFISKRPRTMAAATLKAFKCKQKQVQFYVTHNFSITSS 324  
 DB 304 -FCRICIYLSAINPVIYNLSQKFR---AAFRKLCNCKQKQKPEKPNYSVALSYS 355

## RESULT 14

ID TRFR SHEEP STANDARD; PRT; 398 AA.  
 AC Q28596;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE Thyrotropin-releasing hormone receptor (TRH-R) (Thyroliberin receptor).  
 GN Name=TRHR;  
 OS Ovis aries (Sheep).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Caprinae; Ovis.  
 OX NCBI\_TaxID=9940;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Pituitary;  
 RC MEDLINE=97200775; PubMed=9048604; DOI=10.1210/en.138.3.1019;  
 RA Bockmann J., Boeckers T.M., Winter C., Wittkowski W., Winterhoff H.,  
 RA Deufel T., Krutz M.R.;  
 RT "Thyrotropin expression in hypophyseal pars tuberalis-specific cells is 3,5',3'-triiodothyronine, thyrotropin-releasing hormone, and pit-1 independent.";  
 RT Endocrinology 138:1019-1028 (1997).  
 RL -!- FUNCTION: Receptor for thyrotropin-releasing hormone. This receptor is mediated by G proteins which activate a phosphatidylinositol-calcium second messenger system.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- SIMILARITY: Belongs to the G-protein coupled receptor 1 family.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; X95285; CAA64606.1; -  
 DR InterPro; IPR000276; GPCR Rhodopsin.  
 DR InterPro; IPR009144; ThyrotropinRHR.  
 DR InterPro; IPR002120; TrhRm\_receptor.  
 DR Pfam; PF00001; 7tm1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSIN.  
 DR PRINTS; PR01846; TRHRFAMILY.  
 DR PROSITE; PS00237; G PROTEIN RECEPTOR F1\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1\_2; 1.  
 KW G-protein coupled receptor; Glycoprotein; Transmembrane.  
 FT DOMAIN 1 28 Extracellular (Potential).  
 FT DOMAIN 29 51 1 (Potential).  
 FT DOMAIN 52 61 Cytoplasmic (Potential).  
 FT DOMAIN 62 83 2 (Potential).  
 FT DOMAIN 84 99 Extracellular (Potential).  
 FT TRANSMEM 100 121 3 (Potential).  
 FT TRANSMEM 122 144 Cytoplasmic (Potential).

FT TRANSMEM 145 168 4 (Potential).  
 FT DOMAIN 169 193 Extracellular (Potential).  
 FT TRANSMEM 194 215 5 (Potential).  
 FT DOMAIN 216 266 Cytoplasmic (Potential).  
 FT TRANSMEM 267 288 6 (Potential).  
 FT DOMAIN 289 296 Extracellular (Potential).  
 FT TRANSMEM 297 319 7 (Potential).  
 FT DOMAIN 320 398 Cytoplasmic (Potential).  
 FT DISULFID 98 179 By similarity.  
 FT CARBOHYD 3 3 N-linked (GlcNAc...) (Potential).  
 FT CARBOHYD 10 10 N-linked (GlcNAc...), (Potential).  
 SQ SEQUENCE 398 AA; 45088 MW; 375A31D3DD2A61A CRC64;  
 Query Match 12.7%; Score 249.5; DB 1; Length 398;  
 Best Local Similarity 23.2%; Pred. No. 1.9e-09;  
 Matches 83; Conservative 68; Mismatches 119; Indels 87; Gaps 13;  
 QY 28 VPVYVYLLC-LGLPANILTVIILSQLVARQKSSYN-YLLALAAADILVFFVDF 85  
 DB 26 VVTILLVLIICGLGIVGNIMVVLVVMR--TKHRTPTNCYLVSLAVADLMVL----- 75  
 QY 86 LLEDFILNMQMPQVDDKI-----LEVLFPSSIHSTIWIPTLTDRIYVCH 132  
 DB 76 -----VAAGLPNITDSIYGSWVYGVGCLCTIYQLGINASSCSITAFIERYIAICH 129  
 QY 133 PLKHYTVSPARTKRVISVYITCFLTSIPYYW-----PNIWT-EDYISTSVHVLWI 186  
 DB 130 PIKAQPLCTFSRAKKIIFVMA--FTSIYCLMFLDLNISTYKDALVWSGGYKISR 186  
 QY 187 HCFTVYLVPSCFFILNSIIVKL-----RRKSN----- 215  
 DB 187 YSPYILMDFGVYVPMILATVYGFIAIILFLNIPSDPKNSNMWKNDSHTQKNLN 246  
 QY 216 -----FRLRGYSTGKTTAILFTITSIPATLWAP--RIIMLYHYLYGAPIQNRWLHIMS 267  
 DB 247 SKTSNRYFNSTVSSRKQVTKLAVVILFALLWMPVTLVVNSFLSPFQENWELL--- 303  
 QY 268 DIANMLALLNTAINFLYCFISKRPRTMAAATLKAFKCKQKQVQFYVTHNFSITSS 324  
 DB 304 -FCRICIYLSAINPVIYNLSQKFR---AAFRKLCNCKQKQKPEKPNYSVALSYS 355

## RESULT 15

QY7PYB7 PRELIMINARY; PRT; 440 AA.  
 AC Q7PYB7;  
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)  
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)  
 DE AGCP12601.  
 GN Name=agCGS3608; ORFName=ENSANGG00000018385;  
 OS Anopheles gambiae str. PEST.  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae; Anopheles.  
 OX NCBI\_TaxID=180454;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=PEST;  
 RA Anopheles Genome Sequencing Consortium;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 CC -!- CAUTION: The sequence shown here is derived from an EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is preliminary data.  
 CC EMBL; AA0801008987; EAA01756.1; -  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0001584; P:rhodopsin-like receptor activity; IEA.  
 DR GO; GO:0007186; P:G-protein coupled receptor protein signalin.; IEA.  
 DR InterPro; IPR000276; GPCR\_Rhodopsin.  
 DR Pfam; PF00001; 7tm1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSIN.  
 DR PROSITE; PS00262; G-PROTEIN RECEPTOR F1\_2; 1.  
 SQ SEQUENCE 440 AA; 50079 MW; 959A62F3B6ABC95 CRC64;



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 8, 2005, 18:20:18 ; Search time 129 Seconds  
(without alignments)  
939.235 Million cell updates/sec

Title: US-10-689-832-20

Perfect score: 1967

Sequence: 1 MEHTHAHLAANSLWSWSPG.....NDSKSYQDAIGACVIL 372

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	1967	100.0	372	10	US-09-813-432-20 Sequence 20, Appl
2	1967	100.0	372	14	US-10-219-834-19 Sequence 19, Appl
3	1967	100.0	372	14	US-10-314-076-2 Sequence 2, Appl
4	1967	100.0	372	15	US-10-174-364-20 Sequence 20, Appl
5	1967	100.0	372	15	US-10-333-946-6 Sequence 6, Appl
6	1967	100.0	372	15	US-10-246-583-20 Sequence 20, Appl
7	1967	100.0	372	16	US-10-689-832-20 Sequence 20, Appl
8	1939.5	98.6	369	14	US-10-314-076-4 Sequence 4, Appl
9	1857	94.4	350	10	US-09-813-432-60 Sequence 60, Appl
10	1857	94.4	350	15	US-10-174-364-60 Sequence 60, Appl
11	1857	94.4	350	15	US-10-246-583-60 Sequence 60, Appl
12	1857	94.4	350	16	US-10-689-832-60 Sequence 60, Appl
13	1857	94.4	353	9	US-09-995-225-16 Sequence 16, Appl

14	1857	94.4	353	10	US-09-995-225-16 Sequence 16, Appl
15	1857	94.4	353	14	US-10-094-417-2 Sequence 2, Appl
16	1857	94.4	353	15	US-10-296-294A-5 Sequence 5, Appl
17	1857	94.4	353	15	US-10-450-590-2 Sequence 2, Appl
18	1857	94.4	353	16	US-10-779-104-2 Sequence 2, Appl
19	1853	94.2	350	10	US-09-813-432-59 Sequence 59, Appl
20	1853	94.2	350	15	US-10-174-364-59 Sequence 59, Appl
21	1853	94.2	350	15	US-10-246-583-59 Sequence 59, Appl
22	1853	94.2	350	16	US-10-689-832-59 Sequence 59, Appl
23	1853	94.2	353	10	US-09-813-432-22 Sequence 22, Appl
24	1853	94.2	353	15	US-10-174-364-22 Sequence 22, Appl
25	1853	94.2	353	15	US-10-246-583-22 Sequence 22, Appl
26	1853	94.2	353	16	US-10-689-832-22 Sequence 22, Appl
27	1829	93.0	353	14	US-10-012-140-11 Sequence 11, Appl
28	1741	88.5	345	14	US-10-094-417-20 Sequence 20, Appl
29	1628	82.8	318	15	US-10-174-364-85 Sequence 85, Appl
30	1628	82.8	318	15	US-10-246-583-85 Sequence 85, Appl
31	1623	82.5	333	14	US-10-079-384-26 Sequence 26, Appl
32	1623	82.5	343	15	US-10-450-590-8 Sequence 8, Appl
33	1623	82.5	356	15	US-10-450-590-7 Sequence 7, Appl
34	1623	82.5	385	15	US-10-343-650A-22 Sequence 22, Appl
35	1471	74.8	287	10	US-09-791-932-66 Sequence 66, Appl
36	1426	72.5	272	10	US-09-813-432-57 Sequence 57, Appl
37	1426	72.5	272	10	US-09-813-432-58 Sequence 58, Appl
38	1426	72.5	272	15	US-10-174-364-57 Sequence 57, Appl
39	1426	72.5	272	15	US-10-246-583-57 Sequence 57, Appl
40	1426	72.5	272	16	US-10-689-832-57 Sequence 57, Appl
41	1426	72.5	272	16	US-10-689-832-58 Sequence 58, Appl
42	1317	67.0	265	14	US-10-225-567A-587 Sequence 587, App
43	1317	67.0	265	14	US-10-314-076-14 Sequence 14, Appl
44	912.5	46.4	313	10	US-09-791-932-63 Sequence 63, Appl
45	742	37.7	348	15	US-10-328-916-2 Sequence 2, Appl

# ALIGNMENTS

## RESULT 1

US-09-813-432-20  
; Sequence 20, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majmuder, Kamud  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/193,843

```
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 20
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-813-432-20

Query Match      100.0%; Score 1967; DB 10; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.2e-173;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60

QY 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
DB 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180

QY 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240
DB 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALINTAINPFLYCFISKRPTWAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALINTAINPFLYCFISKRPTWAAATL 300

QY 301 KAFFKCKQPVQFYTNHNFSSITSSPWISPAHSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360
DB 301 KAFFKCKQPVQFYTNHNFSSITSSPWISPAHSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360

QY 361 FEDAIGACVIL 372
DB 361 FEDAIGACVIL 372

RESULT 2
US-10-219-834-19
; Sequence 19, Application US/10219834
; Publication No. US20030096751A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: D0191 NP
; CURRENT APPLICATION NUMBER: US/10/219,834
; PRIOR FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 60/313,658
; PRIOR FILING DATE: 2001-08-20
; PRIOR APPLICATION NUMBER: US 60/340,703
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: US 60/318,675
; PRIOR FILING DATE: 2001-09-12
; PRIOR APPLICATION NUMBER: US 60/355,596
; PRIOR FILING DATE: 2002-02-06
; PRIOR APPLICATION NUMBER: US 60/333,417
; PRIOR FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: US 60/338,367
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 192
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-219-834-19

Query Match      100.0%; Score 1967; DB 14; Length 372;
```

```
Best Local Similarity 100.0%; Pred. No. 3.2e-173;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60

QY 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
DB 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180

QY 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240
DB 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALINTAINPFLYCFISKRPTWAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRMLVHIMSDIANMLALINTAINPFLYCFISKRPTWAAATL 300

QY 301 KAFFKCKQPVQFYTNHNFSSITSSPWISPAHSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360
DB 301 KAFFKCKQPVQFYTNHNFSSITSSPWISPAHSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360

QY 361 FEDAIGACVIL 372
DB 361 FEDAIGACVIL 372

RESULT 3
US-10-314-076-2
; Sequence 2, Application US/10314076
; Publication No. US20030152977A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY34, AND VARIANTS THEREOF
; FILE REFERENCE: D0197NP
; CURRENT APPLICATION NUMBER: US/10/314,076
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: U.S. 60/338,371
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 372
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-314-076-2

Query Match      100.0%; Score 1967; DB 14; Length 372;
Best Local Similarity 100.0%; Pred. No. 3.2e-173;
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60
DB 1 MEHTHAHLAANSSLSWSPGSACGLGFVPVYVYSLLCGLPANILTVIILSQLVARROK 60

QY 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120
DB 61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIEVLEFSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSYSPARTKRVIVSVYITCFLTSIPYYWPNWTDYISTSVH 180

QY 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240
DB 181 HVLIWHCFVTVLVPCSIFFILNSIIIVYKLRRKSNFRLRGYSTGKTTAILFTITSIFATL 240
```



QY 241 WAPRIIMILYHLYGAPIONRWLHVHMSDIANMLALNTAINPFYCFISKRFTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIONRWLHVHMSDIANMLALNTAINPFYCFISKRFTMAAATL 300  
QY 301 KAFFKCKQKQVQFYTHNFSSPWI SPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
Db 301 KAFFKCKQKQVQFYTHNFSSPWI SPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
QY 361 FEDAIGACVIL 372  
Db 361 FEDAIGACVIL 372

RESULT 4  
US-10-174-364-20  
; Sequence 20, Application US/10174364  
; Publication No. US20030216308A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-729C1P2  
; CURRENT APPLICATION NUMBER: US/10/174,364  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836  
; PRIOR FILING DATE: 2000-03-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-174-364-20

Query Match 100.0%; Score 1967; DB 15; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.2e-173;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGACGLGFPVVPVYSLLLCLGLPANILTVILSOLVARROK 60  
Db 1 MEHTHAHLAANSLSWSPGACGLGFPVVPVYSLLLCLGLPANILTVILSOLVARROK 60  
QY 61 SSNYLLALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIIEVLEFSSIHSTWITV 120  
Db 61 SSNYLLALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIIEVLEFSSIHSTWITV 120  
QY 121 PLTIDRYIACHPLKHYTVSPARTKVIIVSVYITCFLTSIPYWWPNWTEYISTSVH 180  
Db 121 PLTIDRYIACHPLKHYTVSPARTKVIIVSVYITCFLTSIPYWWPNWTEYISTSVH 180  
QY 181 HVLWIHCFVTVLVPSCIPFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTTISPAATL 240  
Db 181 HVLWIHCFVTVLVPSCIPFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTTISPAATL 240  
QY 241 WAPRIIMILYHLYGAPIONRWLHVHMSDIANMLALNTAINPFYCFISKRFTMAAATL 300

Db 241 WAPRIIMILYHLYGAPIONRWLHVHMSDIANMLALNTAINPFYCFISKRFTMAAATL 300  
QY 301 KAFFKCKQKQVQFYTHNFSSPWI SPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
Db 301 KAFFKCKQKQVQFYTHNFSSPWI SPANSHCIKMLVYQYDKNGKPKIKSRNDSKSSYQ 360  
QY 361 FEDAIGACVIL 372  
Db 361 FEDAIGACVIL 372

RESULT 5  
US-10-333-946-6  
; Sequence 6, Application US/10333946  
; Publication No. US20040023252A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.  
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.  
; APPLICANT: BURFORD, Neil; YUE, Henry  
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.  
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.  
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.  
; APPLICANT: HAPALIA, April J.A.; YAO, Monique G.  
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.  
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam  
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.  
; APPLICANT: LEE, Ernestine A.; DING, Li  
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: PI-0176 USN  
; CURRENT APPLICATION NUMBER: US/10/333,946  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: PCT/US01/23433  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 60/221,478  
; PRIOR FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 60/223,268  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/227,054  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/231,121  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: US 60/232,243  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 60/232,691  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/235,146  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PERL Program  
; SEQ ID NO 6  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474767CD1  
US-10-333-946-6

Query Match 100.0%; Score 1967; DB 15; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.2e-173;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGACGLGFPVVPVYSLLLCLGLPANILTVILSOLVARROK 60  
Db 1 MEHTHAHLAANSLSWSPGACGLGFPVVPVYSLLLCLGLPANILTVILSOLVARROK 60  
QY 61 SSNYLLALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIIEVLEFSSIHSTWITV 120  
Db 61 SSNYLLALAAADILVLFVDFVDFLEDFILNMQMPQVDPDKIIEVLEFSSIHSTWITV 120  
QY 121 PLTIDRYIACHPLKHYTVSPARTKVIIVSVYITCFLTSIPYWWPNWTEYISTSVH 180

Db 121 PLTIDRYIAVCHPLKYHTVSPARTRKVIVSVVITCFLTSIPYYWPNWTEDYSTSVH 180  
QY 181 HVLWIHCFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
Db 181 HVLWIHCFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
QY 241 WAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALNTAINFFLYCFISKRFRTWAAATL 300  
Db 241 WAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALNTAINFFLYCFISKRFRTWAAATL 300  
QY 301 KAFFKCKQKQVQFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIKSRNDSKSSVQ 360  
Db 301 KAFFKCKQKQVQFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIKSRNDSKSSVQ 360  
QY 361 FEDAIGACVIL 372  
Db 361 FEDAIGACVIL 372

## RESULT 6

US-10-246-583-20  
; Sequence 20, Application US/10246583  
; Publication No. US20040058862A1  
; GENERAL INFORMATION:  
; APPLICANT: Majumder  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-729CIP2CON1  
; CURRENT APPLICATION NUMBER: US/10/246,583  
; PRIOR FILING DATE: 2002-12-06  
; PRIOR APPLICATION NUMBER: 10/174,364  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-246-583-20

Query Match 100.0%; Score 1967; DB 15; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.2e-173;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGSCAGLGFVPVYVYSLLLCGLPANILTVIILSQLVARROK 60  
Db 1 MEHTHAHLAANSLSWSPGSCAGLGFVPVYVYSLLLCGLPANILTVIILSQLVARROK 60  
QY 61 SSYNVLLAALAADILVLFVDFVDFLEDFILNMOMPQVPDKIEVLBFSSHTSIWTV 120  
Db 61 SSYNVLLAALAADILVLFVDFVDFLEDFILNMOMPQVPDKIEVLBFSSHTSIWTV 120  
QY 121 PLTIDRYIAVCHPLKYHTVSPARTRKVIVSVVITCFLTSIPYYWPNWTEDYSTSVH 180  
Db 121 PLTIDRYIAVCHPLKYHTVSPARTRKVIVSVVITCFLTSIPYYWPNWTEDYSTSVH 180

QY 181 HVLWIHCFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
Db 181 HVLWIHCFTVYLVPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240  
QY 241 WAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALNTAINFFLYCFISKRFRTWAAATL 300  
Db 241 WAPRIIMILYHLYGAPIQNRWLHIMSDIANMLALNTAINFFLYCFISKRFRTWAAATL 300  
QY 301 KAFFKCKQKQVQFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIKSRNDSKSSVQ 360  
Db 301 KAFFKCKQKQVQFYTNHNSITSSPWISPAHSHCIKMLVYQYDKNGKPIKSRNDSKSSVQ 360  
QY 361 FEDAIGACVIL 372  
Db 361 FEDAIGACVIL 372

## RESULT 7

US-10-689-832-20  
; Sequence 20, Application US/10689832  
; Publication No. US20040121380A1  
; GENERAL INFORMATION:  
; APPLICANT: Majumder, Kamud  
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729DIV1  
; CURRENT APPLICATION NUMBER: US/10/689,832  
; PRIOR FILING DATE: 2003-10-20  
; PRIOR APPLICATION NUMBER: 09/813,432  
; PRIOR FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: Patent in Ver. 2.1  
; SEQ ID NO 20  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-689-832-20

Query Match 100.0%; Score 1967; DB 16; Length 372;  
Best Local Similarity 100.0%; Pred. No. 3.2e-173;  
Matches 372; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSLSWSPGSCAGLGFVPVYVYSLLLCGLPANILTVIILSQLVARROK 60  
Db 1 MEHTHAHLAANSLSWSPGSCAGLGFVPVYVYSLLLCGLPANILTVIILSQLVARROK 60  
QY 61 SSYNVLLAALAADILVLFVDFVDFLEDFILNMOMPQVPDKIEVLBFSSHTSIWTV 120  
Db 61 SSYNVLLAALAADILVLFVDFVDFLEDFILNMOMPQVPDKIEVLBFSSHTSIWTV 120  
QY 121 PLTIDRYIAVCHPLKYHTVSPARTRKVIVSVVITCFLTSIPYYWPNWTEDYSTSVH 180  
Db 121 PLTIDRYIAVCHPLKYHTVSPARTRKVIVSVVITCFLTSIPYYWPNWTEDYSTSVH 180

QY 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
Db 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
QY 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
QY 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
Db 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
QY 361 FEDAIGACVIL 372  
Db 361 FEDAIGACVIL 372

RESULT 8

US-10-314-076-4  
; Sequence 4, Application US/10314076  
; Publication No. US20030152977A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMV34, AND VARIANTS A  
; FILE REFERENCE: D0197NP  
; CURRENT APPLICATION NUMBER: US/10/314,076  
; PRIOR FILING DATE: 2002-12-06  
; PRIOR APPLICATION NUMBER: U.S. 60/338,371  
; PRIOR FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 4  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-314-076-4

Query Match 98.6%; Score 1939.5; DB 14; Length 369;  
Best Local Similarity 99.2%; Pred. No. 1.1e-170;  
Matches 369; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARROK 60  
Db 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARROK 60  
QY 61 SSNNYLLAALAAADILVLPFIIVFDLLEDFILNMQMPQVDPDKIIEVLEFSSHTSIWTV 120  
Db 61 SSNNYLLAALAAADILVLPFIIVFDLLEDFILNMQMPQVDPDKIIEVLEFSSHTSIWTV 120  
QY 121 PLTIDRYIAVCHPLKHTVSYPARTKVIIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180  
Db 121 PLTIDRYIAVCHPLKHTVSYPARTKVIIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180  
QY 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
Db 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
QY 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
QY 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 360  
Db 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIKSRNDSKSSYQ 357  
QY 361 FEDAIGACVIL 372  
Db 358 FEDAIGACVIL 369

RESULT 9

US-09-813-432-60

; Sequence 60, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majumder, Kamud  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/193,843  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-813-432-60

Query Match 94.4%; Score 1857; DB 10; Length 350;  
Best Local Similarity 100.0%; Pred. No. 4.4e-163;  
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARROK 60  
Db 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARROK 60  
QY 61 SSNNYLLAALAAADILVLPFIIVFDLLEDFILNMQMPQVDPDKIIEVLEFSSHTSIWTV 120  
Db 61 SSNNYLLAALAAADILVLPFIIVFDLLEDFILNMQMPQVDPDKIIEVLEFSSHTSIWTV 120  
QY 121 PLTIDRYIAVCHPLKHTVSYPARTKVIIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180  
Db 121 PLTIDRYIAVCHPLKHTVSYPARTKVIIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180  
QY 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
Db 181 HVLIIHCFVTVLPCSIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTTISIFATL 240  
QY 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIONRWLHIMSDIANLALNTAINFELYCFISKRPRTMAAATL 300  
QY 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIK 350  
Db 301 KAFFKCKQKQPQVYTNHNFSTSSPWISPANSHCIKMLVYQYDKNGKPIK 350

RESULT 10

US-10-174-364-60  
; Sequence 60, Application US/10174364  
; Publication No. US20030216308A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson et al.  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-729CIP2  
; CURRENT APPLICATION NUMBER: US/10/174,364  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836  
; PRIOR FILING DATE: 2000-03-29  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-174-364-60

Query Match 94.4%; Score 1857; DB 15; Length 350;  
Best Local Similarity 100.0%; Pred. No. 4.4e-163;  
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARROK 60  
Db 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARROK 60

QY 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIEVLEPSSHTSIWTV 120  
Db 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIEVLEPSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKHYTVSPARTKVIIVSVYITCFLTSIPYVWPNWTEYISTSVH 180  
Db 121 PLTIDRYIAVCHPLKHYTVSPARTKVIIVSVYITCFLTSIPYVWPNWTEYISTSVH 180

QY 181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLARKSNFRLRGYSTGKTTALLFTTISIFATL 240  
Db 181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLARKSNFRLRGYSTGKTTALLFTTISIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRLVHMSDIANMLALNTAINPFLYCFISKRFRTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIQNRLVHMSDIANMLALNTAINPFLYCFISKRFRTMAAATL 300

QY 301 KAFFKCKQKQVQFYTNHNSITSSPWIISPANSHCIKMLVYQDKNGKPK 350  
Db 301 KAFFKCKQKQVQFYTNHNSITSSPWIISPANSHCIKMLVYQDKNGKPK 350

RESULT 11  
US-10-246-583-60  
; Sequence 60, Application US/10246583  
; Publication No. US20040058862A1  
; GENERAL INFORMATION:  
; APPLICANT: Majumder

; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-729CIP2COM1  
; CURRENT APPLICATION NUMBER: US/10/246,583  
; CURRENT FILING DATE: 2002-12-06  
; PRIOR APPLICATION NUMBER: 10/174,364  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 60  
; LENGTH: 350  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-246-583-60

Query Match 94.4%; Score 1857; DB 15; Length 350;  
Best Local Similarity 100.0%; Pred. No. 4.4e-163;  
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARROK 60  
Db 1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARROK 60

QY 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIEVLEPSSHTSIWTV 120  
Db 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIEVLEPSSHTSIWTV 120

QY 121 PLTIDRYIAVCHPLKHYTVSPARTKVIIVSVYITCFLTSIPYVWPNWTEYISTSVH 180  
Db 121 PLTIDRYIAVCHPLKHYTVSPARTKVIIVSVYITCFLTSIPYVWPNWTEYISTSVH 180

QY 181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLARKSNFRLRGYSTGKTTALLFTTISIFATL 240  
Db 181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLARKSNFRLRGYSTGKTTALLFTTISIFATL 240

QY 241 WAPRIIMILYHLYGAPIQNRLVHMSDIANMLALNTAINPFLYCFISKRFRTMAAATL 300  
Db 241 WAPRIIMILYHLYGAPIQNRLVHMSDIANMLALNTAINPFLYCFISKRFRTMAAATL 300

QY 301 KAFFKCKQKQVQFYTNHNSITSSPWIISPANSHCIKMLVYQDKNGKPK 350  
Db 301 KAFFKCKQKQVQFYTNHNSITSSPWIISPANSHCIKMLVYQDKNGKPK 350

RESULT 12  
US-10-689-832-60  
; Sequence 60, Application US/10689832  
; Publication No. US20040121380A1  
; GENERAL INFORMATION:  
; APPLICANT: Majumder, Kamud  
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729DIV1  
; CURRENT APPLICATION NUMBER: US/10/689,832  
; CURRENT FILING DATE: 2003-10-20  
; PRIOR APPLICATION NUMBER: 09/813,432

; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 60
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-689-832-60

Query Match 94.4%; Score 1857; DB 16; Length 350;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARRQK 60
DB 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARRQK 60
QY 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSHTSIWTV 120
DB 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSHTSIWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSYPARTKRVIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSYPARTKRVIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180
QY 181 HVLWIHCFTVYLVPSCIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
DB 181 HVLWIHCFTVYLVPSCIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFYCFISKRFRTMAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFYCFISKRFRTMAAATL 300
QY 301 KAFFKCKQKQVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350
DB 301 KAFFKCKQKQVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350

RESULT 13
US-09-995-225-16
; Sequence 16, Application US/09995225
; Publication No. US20020193584A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20020193584A1-Endogenous Versions of Human
; FILE REFERENCE: AREN-0308
; CURRENT APPLICATION NUMBER: US/09/995,225
; CURRENT FILING DATE: 2001-11-26
; PRIOR APPLICATION NUMBER: 09/170,496

; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: PCT/US99/23938
; PRIOR FILING DATE: 1998-10-13
; PRIOR APPLICATION NUMBER: 60/253,404
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/255,366
; PRIOR FILING DATE: 2000-12-12
; PRIOR APPLICATION NUMBER: 60/270,286
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/282,365
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/270,266
; PRIOR FILING DATE: 2001-02-20
; PRIOR APPLICATION NUMBER: 60/282,032
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,358
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,356
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/290,917
; PRIOR FILING DATE: 2001-05-14
; PRIOR APPLICATION NUMBER: 60/309,208
; PRIOR FILING DATE: 2001-07-31
; NUMBER OF SEQ ID NOS: 67
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 16
; LENGTH: 353
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: No. US20020193584A1e1 Sequence
; US-09-995-225-16

Query Match 94.4%; Score 1857; DB 9; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARRQK 60
DB 1 MEHTHAHAANSSLSWSPGACGLGFVPVYVYSLLLCLGLPANILTVIILSOLVARRQK 60
QY 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSHTSIWTV 120
DB 61 SSNYLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIEVLEFSSHTSIWTV 120
QY 121 PLTIDRYIAVCHPLKHTVSYPARTKRVIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180
DB 121 PLTIDRYIAVCHPLKHTVSYPARTKRVIVSVYITCFLTSIPYVWPNWTEDYISTSVH 180
QY 181 HVLWIHCFTVYLVPSCIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
DB 181 HVLWIHCFTVYLVPSCIFFILNSIIYVKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
QY 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFYCFISKRFRTMAAATL 300
DB 241 WAPRIIMILYHLYGAPIQNRWLHMSDIANMLALNTAINPFYCFISKRFRTMAAATL 300
QY 301 KAFFKCKQKQVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350
DB 301 KAFFKCKQKQVQFYTNHNSITSSPWISPANSHCIKMLVYQYDKNGKPIK 350

RESULT 14
US-09-995-225-16
; Sequence 16, Application US/09995225
; Publication No. US20030139588A9
; GENERAL INFORMATION:
; APPLICANT: Chen, Ruoping
; APPLICANT: Chu, Zhi Liang
; APPLICANT: Dang, Huang T.
; APPLICANT: Lowitz, Kevin P.
; APPLICANT: Pride, Cameron
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human

```
/ TITLE OF INVENTION: Receptors
/ FILE REFERENCE: AREN-0308
/ CURRENT APPLICATION NUMBER: US/09/995,225
/ CURRENT FILING DATE: 2001-11-26
/ PRIOR APPLICATION NUMBER: 09/170,496
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: PCT/US99/23938
/ PRIOR FILING DATE: 1998-10-13
/ PRIOR APPLICATION NUMBER: 60/253,404
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/255,366
/ PRIOR FILING DATE: 2000-12-12
/ PRIOR APPLICATION NUMBER: 60/270,286
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,365
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/270,266
/ PRIOR FILING DATE: 2001-02-20
/ PRIOR APPLICATION NUMBER: 60/282,032
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,358
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/282,356
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: 60/290,917
/ PRIOR FILING DATE: 2001-05-14
/ PRIOR APPLICATION NUMBER: 60/309,208
/ PRIOR FILING DATE: 2001-07-31
/ NUMBER OF SEQ ID NOS: 67
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 16
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: No. US20030139588A9e1 Sequence
US-09-995-225-16
```

```
Query Match          94.4%; Score 1857; DB 10; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60
Db      1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60

Qy      61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTIWTIV 120
Db      61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTIWTIV 120

Qy      121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLTSIPYVWPNWTEDYISTSVH 180
Db      121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLTSIPYVWPNWTEDYISTSVH 180

Qy      181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
Db      181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240

Qy      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRPRTMAAATL 300
Db      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRPRTMAAATL 300

Qy      301 KAFFKCKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPIK 350
Db      301 KAFFKCKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPIK 350
```

RESULT 15

```
US-10-094-417-2
/ Sequence 2, Application US/10094417
/ Publication No. US20030045685A1
/ GENERAL INFORMATION:
/ APPLICANT: Tian, Hui
```

```
/ APPLICANT: Zhao, Jiagang
/ APPLICANT: Chen, Jin-Long
/ APPLICANT: Cutler, Gene
/ APPLICANT: Tularik Inc.
/ TITLE OF INVENTION: No. US20030045685A1e1 Receptors
/ FILE REFERENCE: 018781-008110US
/ CURRENT APPLICATION NUMBER: US/10/094,417
/ CURRENT FILING DATE: 2002-06-10
/ PRIOR APPLICATION NUMBER: US 09/802,803
/ PRIOR FILING DATE: 2001-03-09
/ PRIOR APPLICATION NUMBER: US 60/276,649
/ PRIOR FILING DATE: 2001-03-16
/ NUMBER OF SEQ ID NOS: 43
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 2
/ LENGTH: 353
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20
US-10-094-417-2
```

```
Query Match          94.4%; Score 1857; DB 14; Length 353;
Best Local Similarity 100.0%; Pred. No. 4.4e-163;
Matches 350; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60
Db      1 MEHTHAHLAANSSLSWSPGSGACGLGFVPVYVYSLLLCLGLPANILTVIILSQLVARQK 60

Qy      61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTIWTIV 120
Db      61 SSYNVLLAALAAADILVLFVDFVDFLEDFILNMQMPQVPDKIIIEVLEFSSIHSTIWTIV 120

Qy      121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLTSIPYVWPNWTEDYISTSVH 180
Db      121 PLTIDRYIAVCHPLKXHTVSYPARTKRVISVYITCFLTSIPYVWPNWTEDYISTSVH 180

Qy      181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240
Db      181 HVLWIHCFTVYLVPCSIFFILNSIIVYKLRKSNFRLRGYSTGKTTAILFTITSIFATL 240

Qy      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRPRTMAAATL 300
Db      241 WAPRIIMILYHLYGAPIQNRWLVIHMSDIANMLALNTAINFFLYCFISKRPRTMAAATL 300

Qy      301 KAFFKCKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPIK 350
Db      301 KAFFKCKQKQPVQFYTNHNFSTSSPWSIPANSHCIKMLVYQYDKNGKPIK 350
```

Search completed: February 8, 2005, 18:31:28  
Job time : 131 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model  
Run on: February 12, 2005, 13:28:42 ; Search time 588 Seconds  
(without alignments)  
3732.944 Million cell updates/sec  
Title: US-10-689-832-20  
Perfect score: 1967  
Sequence: 1 MEHTAHLAANSLSWSPG.....NDKSSYQFIDAIGACVIL 372

Scoring table: BLOSUM62  
Xgapop 10.0, Xgapext 0.5  
Ygapop 10.0, Ygapext 0.5  
Fgapop 6.0, Fgapext 7.0  
Delop 6.0, Delext 7.0

Searched: 5378673 seqs, 2950229984 residues  
Total number of hits satisfying chosen parameters: 10757346  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:  
-MODEL=frame+p2n.model -DEV=xlp  
-Q=/cgn2\_1/USFTO.spool\_p/US10689832/runat\_08022005\_112210\_23265/app.query.fasta\_1.519  
-DB=Published Applications NA -QWFT=fastap -SUFFIX=rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62  
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=2000000000 -USER=US10689832 @CGN 1 1 723 @runat\_08022005\_112210\_23265  
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCX=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
23: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
24: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
25: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
26: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
27: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
28: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
29: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
30: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
31: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
32: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
33: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
34: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
35: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
36: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
37: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
38: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
39: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
40: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
41: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
42: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
43: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
44: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
45: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1967	100.0	1119	10	US-09-813-432-19	Sequence 19, Appl
2	1967	100.0	1119	17	US-10-174-364-19	Sequence 19, Appl
3	1967	100.0	1119	17	US-10-246-583-19	Sequence 19, Appl
4	1967	100.0	1119	18	US-10-689-832-19	Sequence 19, Appl
5	1967	100.0	1130	17	US-10-333-946-25	Sequence 25, Appl
6	1967	100.0	1218	16	US-10-314-076-1	Sequence 1, Appl
7	1939.5	98.6	1110	16	US-10-314-076-3	Sequence 3, Appl
8	1939.5	98.6	2189	14	US-10-219-834-6	Sequence 6, Appl
9	1907	96.9	2117	18	US-10-779-104-1	Sequence 1, Appl
10	1903	96.7	1343	10	US-09-813-432-21	Sequence 21, Appl
11	1903	96.7	1343	17	US-10-174-364-21	Sequence 21, Appl
12	1903	96.7	1343	17	US-10-246-583-21	Sequence 21, Appl
13	1903	96.7	1343	18	US-10-689-832-21	Sequence 21, Appl
14	1879	95.5	1526	14	US-10-012-140-10	Sequence 10, Appl
15	1857	94.4	1059	17	US-10-450-590-1	Sequence 1, Appl
16	1857	94.4	1062	9	US-09-995-225-15	Sequence 15, Appl
17	1857	94.4	1062	10	US-09-995-225-15	Sequence 15, Appl
18	1857	94.4	1062	17	US-10-296-294A-3	Sequence 3, Appl
19	1857	94.4	1062	17	US-10-296-294A-4	Sequence 4, Appl
20	1857	94.4	1202	14	US-10-094-417-1	Sequence 1, Appl
21	1829	93.0	1062	14	US-10-012-140-12	Sequence 12, Appl
22	1741	88.5	1038	14	US-10-094-417-19	Sequence 19, Appl
23	1732	88.1	1062	17	US-10-450-590-9	Sequence 9, Appl
24	1673	85.1	1826	17	US-10-450-590-6	Sequence 6, Appl
25	1628	82.8	957	17	US-10-174-364-84	Sequence 84, Appl
26	1628	82.8	957	17	US-10-246-583-84	Sequence 84, Appl
27	1623	82.5	1002	15	US-10-079-384-25	Sequence 25, Appl
28	1623	82.5	1032	17	US-10-450-590-5	Sequence 5, Appl
29	1623	82.5	1070	17	US-10-450-590-4	Sequence 4, Appl
30	1623	82.5	1158	17	US-10-343-650A-21	Sequence 21, Appl
31	1471	74.8	864	10	US-09-791-932-6	Sequence 6, Appl
32	1317	67.0	795	15	US-10-225-567A-586	Sequence 586, App
33	912	46.4	930	10	US-09-791-932-3	Sequence 3, Appl
34	742	37.7	1466	17	US-10-328-916-4	Sequence 4, Appl
35	737	37.5	1125	9	US-09-995-225-5	Sequence 5, Appl
36	737	37.5	1125	10	US-09-995-225-5	Sequence 5, Appl
37	737	37.5	1427	14	US-10-094-417-13	Sequence 13, Appl
38	737	37.5	1530	13	US-10-011-147-1	Sequence 1, Appl
39	737	37.5	1793	17	US-10-328-916-3	Sequence 3, Appl
40	737	37.5	1962	17	US-10-333-946-29	Sequence 29, Appl
41	733	37.3	1125	14	US-10-012-140-15	Sequence 15, Appl
42	733	37.3	1719	14	US-10-012-140-13	Sequence 13, Appl
43	668	34.0	34118	15	US-10-017-161-1071	Sequence 1071, Ap
44	668	34.0	34118	17	US-10-292-798-909	Sequence 909, App
45	248	12.6	1650	18	US-10-489-425-49	Sequence 49, Appl

ALIGNMENTS

RESULT 1  
US-09-813-432-19  
; Sequence 19, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majmuder, Kamud  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972

```
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-432-19

Alignment Scores:
Pred. No.: 2,948-191 Length: 1119
Score: 1967.00 Matches: 372
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-10-689-832-20 (1-372) x US-09-813-432-19 (1-1119)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaHisSerLeuSerTyrTyrSerProGly 20
Db 1 ATGGAGCAGACGACGACCCACCTCGCAGCCACACAGCTGGTGTCTGGTGTCCCGCCG 60

Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuCysLeuGly 40
Db 1 TCGGCTCGCGCTGGGTTCGTGGCCGCTGCTACTACAGCCTCTTGTGTGCTCGCT 120

Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
Db 1 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAA 180

Qy 61 SerSerTyrAsnTyrLeuAlaLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80
Db 1 TCCTCTACAACTATCTCTTGGCACTCGTGTGCGGACATCTTGGTCTCTTTTCATA 240

Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 1 GTGTTGTGGACTTCCTGTGGAGATTTCATCTTGAACATGCATGCCTCAGGTCCCC 300

Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal 120
Db 1 GACAAGATCATAGAGTGGAAATTCATCATCCATCCACCTCCATATGAGTACTGTA 360

Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 1 CGGTAAACATTGACAGGTATATCGTGTCTGCCACCCGCTCAAGTACACAGGTCTCA 420

Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 1 TACCCAGCCCGCACCCGGAAGTCAATTGTAAGTGTATCATCACCTGTCTCTGTGACG 480

Qy 161 IleProTyrTyrTyrTyrProAsnIleTyrThrGluAspTyrIleSerThrSerValHis 180
Db 1 ATCCCCCTATTACTGTGGTGGCCCAACATCTGGACTGAAGACTATACAGCACCTCTGTG 540

Qy 181 HisValLeuIleTyrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 1
```



```
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
US-10-174-364-19

Alignment Scores:
Pred. No.: 2,94e-191 Length: 1119
Score: 1967.00 Matches: 372
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-689-832-20 (1-372) x US-10-174-364-19 (1-1119)

QY 1 MetGluHisThrHisAlaHisLeuAlaAAsenSerLeuSerTrpTrpSerProGly 20
DB 1 ATGGAGCACAGCAGCAGCCACCTCGCAGCCACAGCTCGCTGTCTGGTGGTCCCGGC 60
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrSerLeuLeuLeuGly 40
DB 61 TGGCCTCGCGCTTGGGTTTGGGCCCGGGTCTACTACAGCCTCTGTGTGCTCGGT 120
QY 41 LeuProAlaAsnIleLeuThrValIleLeuSerGlnLeuValAlaAArgGlnLys 60
DB 121 TTACGACGAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAGAGACAGAG 180
QY 61 SerSerTyAsnTyrLeuLeuAlaAlaAAspIleLeuValLeuPhePheIle 80
DB 181 TCTCTTACAATATCTCTGGCAGCTGGTGGTGGCAGCATCTGGTCTCTTTTTCATA 240
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
DB 241 GTGTTGTGGATCTCTGTGTGAAGATTTTCATCTTGAAATGAGATGCTCAGGTCC 300
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
DB 301 GACAGATCATAGAGTGTGGAAATCTCATCTCATCCACCTCCATGATGATTAAGTGA 360
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
DB 361 CGGTTAAACCATTCACAGGTATATCGCTGTGCTGCCAGCTCAAGTACCACACGGTCTCA 420
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
DB 421 TACCCAGCCCGCACCCGGAAGTCATTGTAAGTGTATTACATCACCTGCTTCTCGACCCAGC 480
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
DB 481 ATCCCTTATATCTGGTGGCCCAACATCTGAGCTGAGACTATCATGACCTCTGTGTCAT 540
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
DB 541 CAGGCTCTCATCTGGATCCACTGCTTACCGCTACTGCTGGTGGTGGTGGTGGTGGTGGT 600
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
DB 601 ATCTTGAATCAATCATCTGTAAGCTCAGGAGGAGGAGCAATTTTGTCTCGGTGGC 660
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
DB 661 TACTCCAGGGGAGACCCGCCATCTTGTTCACCATTTACCTTCCATCTTCCACACTT 720
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValProIleGlnAsnArg 260
DB 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGCGCCCATCCAGAACCCG 780
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
DB 781 TGGTGGTACATCATGATTCGACATTTGCCAATGTCATGAGCTTAGCCCTTCTGAACACAGCCATC 840

; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP3CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURES:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
US-10-246-583-19

Alignment Scores:
Pred. No.: 2,94e-191 Length: 1119
Score: 1967.00 Matches: 372
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

US-10-689-832-20 (1-372) x US-10-246-583-19 (1-1119)

QY 1 MetGluHisThrHisAlaHisLeuAlaAAsenSerLeuSerTrpTrpSerProGly 20
```

```
Db      1  ATGGAGCACACGCGCCACCTCGACGCCAACAGCTCGTGTCTTGGTGTCCCGGCG 60
QY      21  SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db      61  TCGGCTCGGGCTTGGGTTCGTGCCCGTGGTGTACTACAGCTCTTGTCTGTCCGTGGT 120
QY      41  LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
Db      121  TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180
QY      61  SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      181  TCCTCTACACTATCTCTGGCACTCGTGTCCGACATCTTGGTCTCTTTTCATA 240
QY      81  ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241  GTGTTGTGGACTTCTCTGTGAAGATTCTATCTTGAACATGCAGATGCCTCAGGTCCC 300
QY      101  AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db      301  GACAAGATCATAGAAGTCTGAATCTCATCTCCACCTCCATATGATGATCTACTGTA 360
QY      121  ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db      361  CGTTTAACTTACAGAGTATATCGTGTGCCACCCGCTCAAGTACCAACAGGTCTCA 420
QY      141  TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db      421  TACCAGCCGCGACCCGGAAGTCATTGTAGTGTATACATCACCTGCTTCTGACACAGC 480
QY      161  IleProTyrTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db      481  ATCCCTTATTACTGTGTGCCCAACATCTGGACTGAAGACTCATCATCAGCACCTCTGTGCAT 540
QY      181  HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db      541  CAGTCTCTCATCTGGATCCAGTCTTACCCTTACCTGTGGTGGCTGTCTCATCTCTTC 600
QY      201  IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db      601  ATCTTGAACCTAATCATTTGTGTACAACTCAGAGGAAAGCAATTTTCGTCTCGTGGC 660
QY      221  TyrSerThrGlyThrThrAlaIlePheThrIleThrSerIlePheAlaThrLeu 240
Db      661  TACTCCAGGGGAAGACACCGCCATCTGTTCACCATTACCTCCATCTTTGCCACACTT 720
QY      241  TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValProIleGlnAsnArg 260
Db      721  TGGGCCCCCGCATCATCATGATCTTTTACCACCTCTATGGGGCCCATCCAGAACCGC 780
QY      261  TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db      781  TGGCTGGTACACATCATGTCCGACATTTGCCACATGTAGCCCTTCTGAACACAGCCATC 840
QY      281  AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db      841  AACTTCTTCTCTACTGTCTTCATCAGCAAGCGGTTCCGACCATCGGACGCGCCACGCTC 900
QY      301  LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db      901  AAGGCTTTCTTCAAGTGCACCAAGCAACCTGTACAGTTCTACCAATCATNACTTTTCC 960
QY      321  IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db      961  ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACACTGCATCAAGATGTGTGTAC 1020
QY      341  GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerTyrGln 360
Db      1021  CAGTATGACAAAATGGAAAACCTATAAAAAGTCGTAATGACAGCAAAAGCTCTTACCAG 1080
QY      361  PheGluAspAlaIleGlyValAlaCysValIleIleLeu 372
```

```
Db      1081  TTTGAGATGCCATTGGAGCTTGTGTCAATCATCTG 1116
RESULT 4
US-10-689-832-19
; Sequence 19, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmudar, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; PRIOR FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-689-832-19
Alignment Scores:
Pred. No.:      2,94e-191      Length:      1119
Score:          1967.00      Matches:      372
Percent Similarity: 100.00%      Conservative: 0
Best Local Similarity: 100.00%      Mismatches: 0
Query Match:      100.00%      Indels:      0
DB:              18          Gaps:        0
US-10-689-832-20 (1-372) x US-10-689-832-19 (1-1119)
QY      1  MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20
Db      1  ATGGAGCACACGCGCCACCTCGACGCCAACAGCTCGTGTCTTGGTGTCCCGGCG 60
QY      21  SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db      61  TCGGCTCGGGCTTGGGTTCGTGCCCGTGGTGTACTACAGCTCTTGTCTGTCCGTGGT 120
QY      41  LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
Db      121  TTACCAAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180
QY      61  SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db      181  TCCTCTACACTATCTCTGGCACTCGTGTCCGACATCTTGGTCTCTTTTCATA 240
QY      81  ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db      241  GTGTTGTGGACTTCTCTGTGAAGATTCTATCTTGAACATGCAGATGCCTCAGGTCCC 300
QY      101  AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
```

Db 301 GACAGATCATGAGAGTGTGGAAATCTCATCCATCCACACCTCCATATGGATTACTGTA 360  
QY 121 ProLeuThrlleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db 361 CCGTTAACCATTTGACAGGTATATCGTGTCTGCCACCGCTCAAGTACACACGGTCTCA 420  
QY 141 TyrProAlaAargThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 421 TACCCAGCCCGCACCCGGAAGTCAATGTAAAGTGTATTACATCACCTGCTTCTCGACAGC 480  
QY 161 IleProTyrTyrTrpTrpProLeuIleThrThrGluAspTyrIleSerThrSerValHis 180  
Db 481 ATCCCTATTACTGGTGGCCCAACATCTGACCTGAAGACTACATACGACCTCTGTCGAT 540  
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db 541 CAGTCTCTCATCTGGATCCACTGCTTCCCGTCTACCTTGGTGGCTGCTCCATCTCTTC 600  
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 601 ATCTTGAACCTCAATCATTTGTGTACAAGCTCAGGAGGAGCAATTTTCTGCTCCGTGTC 660  
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 661 TACTCACCGGGAAGACACCGCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720  
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
Db 721 TGGGCCCCCGCATCATCATGATCTTTACCACTCTATGGGGGGCCCATCCAGAACCGC 780  
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
Db 781 TGGTGGTACATCATCATGCTCCGACATGTCACATGCTAGCCCTTCTGAACACAGCCATC 840  
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCGCCACCATGGCAGCCGCCAGCTC 900  
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db 901 AAGCTTTCTTCAAGTGCAGAGCAAGCACTGTACAGTTCTACACCATCATCACTTTTCC 960  
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
Db 961 ATAACAAGTAGCCCTCGATCTCGCCGCAAACTCACACTGCATCAAGATGCTGGGTAC 1020  
QY 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerThrGln 360  
Db 1021 CAGTATGACAAAATGGAAAACCTATAAAAAGTCGTAATGACAGCAAAAGCTCTTACCAG 1080  
QY 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
Db 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTG 1116

## RESULT 5

US-10-333-946-25  
; Sequence 25, Application US/10333946  
; Publication No. US20040023252A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.  
; APPLICANT: ARVIZU, Chandra S.; LAL, Preeti G.  
; APPLICANT: BURFORD, Neil; YUE, Henry  
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.  
; APPLICANT: RAMKUMAR, Jayalaxmi; BAUGHN, Mariah R.  
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.  
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.  
; APPLICANT: LU, Yan; TRIBOULET, Catherine M.  
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam  
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.  
; APPLICANT: LEE, Ernestine A.; DING, Li  
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: PI-0176 USN  
; CURRENT APPLICATION NUMBER: US/10/333,946

; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: PCT/US01/23433  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 60/221,478  
; PRIOR FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 60/223,268  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/227,054  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/231,121  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: US 60/232,243  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 60/232,691  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/235,146  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PERL Program  
; SEQ ID NO 25  
; LENGTH: 1130  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474767CBI  
US-10-333-946-25

## Alignment Scores:

Pred. No.: 2,996-191 Length: 1130  
Score: 1967.00 Matches: 372  
Percent Similarity: 100.00% Conservative: 0  
Best Local Similarity: 100.00% Mismatches: 0  
Query Match: 100.00% Indels: 0  
DB: 17 Gaps: 0

US-10-689-832-20 (1-372) x US-10-333-946-25 (1-1130)

QY 1 MetGluHisThrHisAlaHisLeuAlaHisSerSerLeuSerTrpTrpSerProGly 20  
Db 12 ATGGAGCACACGACGCCACCTCCAGCCAAACAGCTGCTGTCTTGGTGGTCCCCGGC 71  
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
Db 72 TCGGCTCGGGCTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 131  
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAargArgGlnLys 60  
Db 132 TTACCAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGACAGAG 191  
QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80  
Db 192 TCCTCTCAACTATCTCTTGGCACTCGCTGCTGGCACTCTTGGTCTCTCTTTTCA 251  
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db 252 GTGTTTGGACTTCTCTGTGGAGATTTTCATCTTGAACATGCAGATGCCTCAGGTCCC 311  
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
Db 312 GACAGAGCATAGAGAGTGTGGAAATTCATCATCCACCATCCATATGGATTATGTA 371  
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
Db 372 CGTTAACCATTTGACAGGTATATCTGCTTGGCCACCCGCTCCAGTACACAGGTCTCA 431  
QY 141 TyrProAlaAargThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 432 TACCCAGCCCGCACCCGGAAGTCAATGTAAAGTGTATTACATCACCTGCTTCTCGACAGC 491  
QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
Db 492 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 551

```
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
|||
Db 552 CAGGICCTCATCTGGATCCATCGCTTACCGGCTACCTGGTGGCCCTGCTCAATCTCTTC 611
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheArgLeuArgGly 220
|||
Db 612 ATCTTGAACCTCAATCATTTGTGTACAACTCAGAGAGAGAGCAATTTTCGTCCTCGTGGC 671
QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
|||
Db 672 TACTCCACGGGAGAGACCACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 731
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisIleTyrGlyAlaProIleGlnAsnArg 260
|||
Db 732 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCGC 791
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
|||
Db 792 TGGCTGGTACACATCATGTCCGACATTTGCCACATGCTAGCCCTTCTGACACAGCCATC 851
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
|||
Db 852 AACTTCTTCTCTACTGCTTCAATCAGCAAGCGGTTCCGCACCATGGCAGCGCCACGCTC 911
QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
|||
Db 912 AAGGCTTTTCTCAAGTGCAGAGCAACCTGTACAGTTCTACCAATCATAAATCTTTTCC 971
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
|||
Db 972 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGATC 1031
QY 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360
|||
Db 1032 CAGTATGACAAATGGAACCTATAAAAGTCGTAATGACAGCAAAAGCTCTCTACCAG 1091
QY 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372
|||
Db 1092 ITTGAAGATGCCATTGGAGCTTGTGTATCATCTCTG 1127

RESULT 6
US-10-314-076-1
; Sequence 1, Application US/10314076
; Publication No. US2003015297A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY34, AND VARIANTS A
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: D0197NP
; CURRENT APPLICATION NUMBER: US/10/314,076
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: U.S. 60/338,371
; PRIOR FILING DATE: 2001-12-06
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2198
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1047)..(2162)
; OTHER INFORMATION:
US-10-314-076-1

Alignment Scores:
Pred. No.: 8,36e-191 Length: 2198
Score: 1967.00 Matches: 372
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0
```

Db 2067 CAGTATGACAAATGGAACCTATAAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 2126  
Qy 361 PheGluAspAlaIleGlyAlaCysValIleLeu 372  
Db 2127 TTTGAAGATGCCATTGGAGCTTGTCATCATCTTG 2162  
RESULT 7  
US-10-314-076-3  
; Sequence 3, Application US/10314076  
; Publication No. US20030152977A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPBMY34, AND VARIANTS A  
; TITLE OF INVENTION: METHODS OF USE THEREOF  
; FILE REFERENCE: D0197NP  
; CURRENT APPLICATION NUMBER: US/10/314,076  
; CURRENT FILING DATE: 2002-12-06  
; PRIOR APPLICATION NUMBER: U.S. 60/338,371  
; PRIOR FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 3  
; LENGTH: 1110  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (1)..(1107)  
; OTHER INFORMATION:  
US-10-314-076-3  
Alignment Scores:  
Pred. No.: 1.91e-188 Length: 1110  
Score: 1939.50 Matches: 369  
Percent Similarity: 99.19% Conservative: 0  
Best Local Similarity: 98.19% Mismatches: 0  
Query Match: 98.60% Indels: 3  
DB: 16 Gaps: 1  
US-10-689-832-20 (1-372) x US-10-314-076-3 (1-1110)  
Qy 1 MetGluHisThrHisAlaHisLeuAlaAAsnSerLeuSerTrpTrpSerProGly 20  
Db 1 ATGGAGCACAGCCAGCCACCTCGCAGCAGCAGCTCGCTGCTTGGTGGTCCCGGC 60  
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyTrpSerLeuLeuLeuGly 40  
Db 61 TCGGCTGGCGTTGGGTTTCGTCGCCGCGTCTACTACAGCTCTTGTCTGTCCTCGT 120  
Qy 41 LeuProAlaAsnIleLeuThrValIleLeuSerGlnLeuValAlaAargGlnLys 60  
Db 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGA 180  
Qy 61 SerSerTyAsnTyLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheLe 80  
Db 181 TCCTCTACAACATATCTCTGGCACTCGCTGTCGCCGACATCTGGTCTCTTTTCATA 240  
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAAsnMetGlnMetProGlnValPro 100  
Db 241 GGTGTGGGATCTCTGTGGAAGATTTCATCTTGAATCATGATGATGCTCAGTCCCG 300  
Qy 101 AspyIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
Db 301 GACAAGATCATAGAAGTGTGGAATCTCATCATCCACCTCCATATGATGATATCTGA 360  
Qy 121 ProLeuThrIleAspArgTyIleAlaValCysHisProLeuTyTrpHisThrValSer 140  
Db 361 CCGTTAACCAATGACAGGTATATCGCTGTCTGCCCGCTCAAGTACCACACGCTCTCA 420  
Qy 141 TyTrpAlaArgThrArgLysValIleValSerValTyIleThrCysPheLeuThrSer 160  
Db 421 TACCAGCCCGCACCCGGAAGTCAATTGTAAGTGTTCATCATCTGCTTCTGACCGAGC 480

Qy 161 IleProTyTrpTrpTrpProAsnIleTrpThrGluAspTyTrpIleSerThrSerValHis 180  
Db 481 ATCCCTATTACTGGTGGCCCAACATCTGGACTGAAGACTATACATCAGCACCTCTGTGCAT 540  
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyLeuValProCysSerIlePhePhe 200  
Db 541 CACGTCCCTCATCTGGATCCATGCTTACCGCTTACCTGCTGGTGGCTTCCATCTTCTTC 600  
Qy 201 IleLeuAAsnSerIleIleValTyTrpLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 601 ATCTTGAATCAATCATCTTGTGTACAGCTCAGGAGAGAGCAATTTTCTCTCCGTGGC 660  
Qy 221 TyTrpSerThrGlyThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 661 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720  
Qy 241 TrpAlaProArgIleIleMetIleLeuTyHisLeuTyGlyAlaProIleGlnAsnArg 260  
Db 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGGGGCGCCCATCCAGAACCCGC 780  
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAAsnMetLeuAlaLeuLeuAAsnThrAlaIle 280  
Db 781 TGGCTGGTACACATCATCTCCGACATTCGCAACATGCTAGCCCTTCTGACACAGCCATC 840  
Qy 281 AsnPhePheLeuTyTrpCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300  
Db 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCCGACCATCGCAGCGCCACGCTC 900  
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyThrAsnHisAsnPheSer 320  
Db 901 AAGGCTTTCTTCAAGTGCAGAGCACTGTACAGTTCTACCAATCATTAACATTTTCC 960  
Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyTr 340  
Db 961 ATAACAAGTAGCCCTCGATCTCGCGGCAAACTCAGACTGATCAGATGCTGGTGTAC 1020  
Qy 341 GlnTyAspLysAsnGlyLysProLysSerArgAsnAspSerLysSerSerTyGln 360  
Db 1021 CAGTATGACAAAAATGGA-----AAAAGTCGTAATGACAGCAAAAGCTCCTACCAG 1071  
Qy 361 PheGluAspAlaIleGlyAlaCysValIleLeu 372  
Db 1072 TTTGAAGATGCCATTGGAGCTTGTCATCATCTG 1107  
RESULT 8  
US-10-219-834-6  
; Sequence 6, Application US/10219834  
; Publication No. US20030096751A1  
; GENERAL INFORMATION:  
; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE TH  
; FILE REFERENCE: D0191 NP  
; CURRENT APPLICATION NUMBER: US/10/219,834  
; CURRENT FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 60/313,658  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: US 60/340,703  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: US 60/318,675  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 60/355,596  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: US 60/333,417  
; PRIOR FILING DATE: 2001-11-26  
; PRIOR APPLICATION NUMBER: US 60/338,367  
; PRIOR FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 2189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-219-834-6

Alignment Scores:  
Pred. No.: 5,478-188 Length: 2189  
Score: 1939.50 Matches: 369  
Percent Similarity: 99.19% Conservative: 0  
Best Local Similarity: 99.19% Mismatches: 0  
Query Match: 98.60% Indels: 3  
DB: 14 Gaps: 1

US-10-689-832-20 (1-372) x US-10-219-834-6 (1-2189)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
DB 1047 ATGGAGCACACGACGCCACCTCGACGACCAACAGCTCGCTGTCTGGTCCCGCCG 1106

QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
DB 1107 TCGGCTCGCGCTTGGGTTTCGTGCCGTGTACTACAGCCTCTTGTCTCGCTCGGT 1166

QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60  
DB 1167 TTACAGCAAAATATCTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 1226

QY 61 SerSerTyrAsnTyrLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
DB 1227 TCCTCTACAACTATCTCTGGCACTCGCTGCGACATCTTGGTCTCTTTTCATA 1286

QY 81 ValPheValAppPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
DB 1287 GTGTTGTGGACTTCCTGTGGAGATTTTCATCTTGAACATGCAGATGCCTCAGGTCCC 1346

QY 101 AspyIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
DB 1347 GACAAGATCATAGAAGTGTGAAATCTCATCCATCCACCTCATATGATGATCTACTGTA 1406

QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140  
DB 1407 CGTTAAACATGTACAGGTATATCGCTGTGCCACCCGCTCAAGTACCACAGGTCTCA 1466

QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
DB 1467 TACCAGCCCGCACCCGGAAGTCAATGTAGTGTTTACATCACCTGCTCTCTGACCAGC 1526

QY 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180  
DB 1527 ATCCCTCTATCTGTGTGGCCCAACATCTGACCTGAAGACTACATCAGCACCTCTGTGAT 1586

QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
DB 1587 CAGGTCTCATCTGGATCCACTGCTTCCCGTCTACCTGGTGGCTGCTCCATCTTCTTC 1646

QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPhaArgLeuArgGly 220  
DB 1647 ATCTTGAATCAATCATTTGTGTACAAAGCTCAGGAGAAAGACAAATTTTCGTCTCG 1706

QY 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
DB 1707 TACTCCAGGGGAACACACCCCACTTGTTCACATTTACCTTCATCTTGGCCACACTT 1766

QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260  
DB 1767 TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGGCCCATCCAGAACCGC 1826

QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
DB 1827 TGGCTGGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTCTTGAACACAGCCATC 1886

QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300  
DB 1887 AACTTCTTCTCTACTGCTTCTATCAGAACGGTTCGGACCATGGACGGCCGACGCTC 1946

QY 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheThrAsnHisAsnPheSer 320

DB 1947 AAGGCTTTCTCAAGTCCAGAGCAACCTGTACAGTTCTACACCAATCATACTTTTC 2006

QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
DB 2007 ATAAACAAGTAGCCCTCGATCTCGCCGCAAACTCACACTGCATCAAGATGCTGTGTAC 2066

QY 341 GlnTyrAspLysAsnGlyLysProIleLysSerArgAsnAspSerLysSerSerTyrGln 360  
DB 2067 CAGTATGACAAATATGA-----AAAGTCTGTAATGACAGCAAAAGCTCTCTACAG 2117

QY 361 PheGluAspAlaIleGlyAlaCysValIleIleLeu 372  
DB 2118 TTGAAGATGCCATTGGAGCTTGTGTATCATCATCCTG 2153

RESULT 9  
US-10-779-104-1  
; Sequence 1, Application US/10779104  
; Publication No. US20040161799A1  
; GENERAL INFORMATION:  
; APPLICANT: Susan Croll-Kalish  
; TITLE OF INVENTION: KOR3like-Proteins and Methods of Modulating KOR3L-Mediated Activi  
; FILE REFERENCE: REG 1000A  
; CURRENT APPLICATION NUMBER: US/10779,104  
; CURRENT FILING DATE: 2004-02-13  
; PRIOR APPLICATION NUMBER: 60/447,447  
; PRIOR FILING DATE: 2003-02-14  
; PRIOR APPLICATION NUMBER: 60/495,577  
; PRIOR FILING DATE: 2003-08-14  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 2117  
; TYPE: DNA  
; ORGANISM: homo sapiens  
US-10-779-104-1

Alignment Scores:  
Pred. No.: 1,11e-184 Length: 2117  
Score: 1907.00 Matches: 372  
Percent Similarity: 83.22% Conservative: 0  
Best Local Similarity: 83.22% Mismatches: 0  
Query Match: 96.95% Indels: 75  
DB: 18 Gaps: 1

US-10-689-832-20 (1-372) x US-10-779-104-1 (1-2117)

QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTrpTrpSerProGly 20  
DB 280 ATGGAGCACACGACGCCACCTCGACGCAACAGCTCGCTGTCTTGGTGGTCCCGCCG 339

QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
DB 340 TCGGCTCGCGCTTGGGTTTCGTGCCGTGTCTTACTACAGCCTCTTGTCTGTGCTCGGT 399

QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60  
DB 400 TTACAGCAAAATATCTTGAACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 459

QY 61 SerSerTyrAsnTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80  
DB 460 TCCTCTTACAACTATCTTGGCACTCGCTGCTGCCGACATCTTGGTCTCTTTTCATA 519

QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
DB 520 GTGTTGTGGACTTCCTGTTGGAAGATTTTCATCTTGAACATGCAGATGCCTCAGGTCCCC 579

QY 101 AspyIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
DB 580 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACCTCATATGATGATCTACTGTA 639

QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140

Db 640 CCGTTAACCAATTGACAGGTATATCGCTGTCTGCCACCCGCTCAAGTACACACGGTCTCA 699  
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160  
Db 700 TACCAGCCCGCACCGGAAAGTCATTGTAAAGTTTATCATCACTGCTTCTTGACACG 759  
Qy 161 IleProTyrTyrTrpTrpProIleThrThrGluAspTyrIleSerThrSerValHis 180  
Db 760 ATCCCTATTACTGTGTGGCCCAACATCTGGAAGTGAAGACTATACAGCACCTCTGTGCAT 819  
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200  
Db 820 CAGGCTCTCATGTGATCCACCTGCTTACCGTCTACCTGTGGTCCCTGTCTCACTCTTC 879  
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220  
Db 880 ATCTTGAAGTCAATCATGTGTACAAAGCTCAGAGGAGAGCAATTTTCGTCTCGTGGC 939  
Qy 221 TyrSerThrGlyLysThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240  
Db 940 TACTCCAGGGGAAGACACACCGCCATCTTGTTCACCAATTTACCTCCATCTTTGCCACATT 999  
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyValProIleGlnAsnArg 260  
Db 1000 TGGGCCCCCGCATCATCATGATTTCTTACCACCTCTATGGGGCGCCCATCCAGAACCGC 1059  
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280  
Db 1060 TGGCTGTGTGCATCATGTCCGACATTTGCCACATGCTAGCCCTTCTGNACACACCATC 1119  
Qy 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrLeu 300  
Db 1120 AACTTCTTCTCTACTGCTTTCATGCAAGCGTTCCGCACCATGGCAGCGCCACGCTC 1179  
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320  
Db 1180 AAGGCTTTTCTCAAGTGGCAGAAAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 1239  
Qy 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340  
Db 1240 ATAACAGTAGGCTTGGATCTCGCGGCAAACTCACACTGCATCAGATGCTGTGTATC 1299  
Qy 341 GlnTyrAspLysAsnGlyLysProIleLys----- 350  
Db 1300 CAGTATGACAAATAGGAAACCTATAAAAGTATCCCGTGATCCCATAGGTGGCAAC 1359  
Qy 350 ----- 350  
Db 1360 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGGTGTCCATCTATGGCTGACGACG 1419  
Qy 350 ----- 350  
Db 1420 TCTCCTTAAGAGTGCTAATCCGATTTCTGTCTCCGCGAGAGTGGCAATTTCTCAGACTG 1479  
Qy 350 ----- 350  
Db 1480 GTAGATGAGAGAGATGGAAGAGAGAAAGAGAGAGCATGAAGCTTGTTTTACTTATGCA 1539  
Qy 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG 366  
Db 1540 TTTATTTCCACAGAGTCGTAATGACAGCAAAAGCTCTACCAAGTTTGAAGATGCCATTGG 1599  
Qy 366 ValaCysValIleIleLeu 372  
Db 1600 AGCTTGTGTATCATCTCTG 1618

RESULT 10

US-09-813-432-21  
; Sequence 21, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majmuder, Kamud

; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/193,843  
; PRIOR FILING DATE: 2000-03-31  
; NUMBER OF SEQ ID NOS: 78  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 1343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-813-432-21

Alignment Scores:  
Pred. No.: 1,41e-184 Length: 1343  
Score: 1903.00 Matches: 271  
Percent Similarity: 83.00% Conservative: 0  
Best Local Similarity: 83.00% Mismatches: 1  
Query Match: 96.75% Indels: 75  
DB: 10 Gaps: 1

US-10-689-832-20 (1-372) x US-09-813-432-21 (1-1343)

Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaHisSerSerLeuSerTrpTrpSerProGly 20  
Db 2 ATGGAGCACACGACGCGCCACCTCGACGCCAACAGCTCGTGTCTTGGTGGTCCCGCGC 61  
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40  
Db 62 TCGGCTCGCGCTTGGGTTTCTGTCCTTACTACAGCTTCTGTGTCTGTCTGTCTGTCT 121  
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgArgGlnLys 60  
Db 122 TTACAGCAATATCTTTCAGAGTGTATCTCTCTCCAGCTGGTGGCAGAGACAGAG 181  
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80  
Db 182 TCCCTCTACAACTATCTCTTGGCACTCGCTGCGGACATCTTGGTCTCTCTTTTTCATA 241  
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100  
Db 242 GTGTTTGTGGACTTCTCTTGGAAAGATTTCATCTTGAACATGCGACATGCTCAGTCC 301  
Qy 101 AspLysValIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120  
Db 302 GACAGAGTATAGAGTGTCTGGAATCTCTATCCATCCACCTCATATGGAATTACTGTA 361  
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140

```
|||||
362 CCGTTAACCATTTGACAGGTATATCACTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 421
QY TyrProAlaAthrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db TACCAGCCCGCACCCGGAAGTCAATTGAAGTGTTTACATCACCTGCTTCTGACCAAGC 481
QY IleProTyrTyrTyrTyrProAsnIleThrPheGluAspTyrIleSerThrSerValHis 180
Db ATCCCTTATTACTGTGTGCCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 541
QY HisValLeuIleThrPheHisCysPheThrValTyrIleuValProCysSerIlePhePhe 200
Db CACGTCCTCATCTGGATCATCTGTTACCGCTTACCTGGTGGCTGCTCCATCTTCTTC 601
QY IleLeuAenSerIleIleValTyrIysLeuArgLysSerAsnPheArgLeuArgGly 220
Db ATCTTGAACCTCAATCATTTGTGTACAGCTCAGGAGGAAGACAAATTTTCGTCCGTGGC 661
QY TyrSerThrGlyThrThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
Db TACTCCAGGGGAAGACACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 721
QY TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db TGGGCCCCCGCATCATCATGATCTTTTACCACTTATGGGGGCCCATCCAGACCGC 781
QY TrpLeuValHisIleMetSerAspIleAlaIleMetLeuAlaLeuLeuAenThrAlaIle 280
Db TGGCTGTGTACATCATCTCGACATTCGCCAATGCTAGCCCTTCTGAACACAGCCATC 841
QY AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrIleu 300
Db AACTTCTTCTCTACTGTCTTATCAGCAAGCGGTTCGCCACATGGCAGCGCCACGCTC 901
QY LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db AAGGCTTCTTCAAGTGCAGAGCAACTGTACAGTCTTACACCAATCATTAATCTTTC 961
QY IleThrSerSerProTyrIleSerProAlaAenSerHisCysIleLysMetLeuValTyr 340
Db ATAAACAAGTAGCCCTGGATCTCGCGCGCAAACTCACACTGCATCAAGATGCTGTGTAC 1021
QY GlnTyrAspLysAsnGlyAspProfile-Lys----- 350
Db CAGTATGACAAAATGGAAAAACCTATAAAAGTATCCCGTGATTCATAGGTGGCAAC 1081
QY 350 ----- 350
Db TACTGCCTCTGTCTAATCCATTTCCAGATGGGAGGTGTCCCATCCTATGCTGAGCAGC 1141
QY 350 ----- 350
Db TCTCCTTAAGAGTGTAAATCCGATTTCTGTCTCCCGCAGACTGGGCAATTCACAGACTG 1201
QY 350 ----- 350
Db GTAGATGAGAAGAGATGGAAGAGAAAGAGAGCATGAAAGCTTTGTTTACTTTATGCA 1261
QY 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366
Db TTTATTTCCACAGAGTGTAAATGACAGCAAAAGCTCTTACCAGTTTGAAGATGCCATTTGG 1321
QY yAlaCysValIleIleLeu 372
Db AGCTTGTGTATCATCTCTG 1340
```

RESULT 11

```
US-10-174-364-21
; Sequence 21, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
```

```
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2)..(1060)
US-10-174-364-21
Alignment Scores:
Pred. No.: 1,41e-184 Length: 1343
Score: 1903.00 Matches: 371
Percent Similarity: 83.00% Conservative: 0
Best Local Similarity: 83.00% Mismatches: 1
Query Match: 96.75% Indels: 75
DB: 17 Gaps: 1
US-10-689-832-20 (1-372) x US-10-174-364-21 (1-1343)
QY 1 MetGluHisThrHisAlaHisLeuAlaAaenSerSerLeuSerTyrTyrSerProGly 20
Db 2 ATGGAGCACACGACGCGCCACCTCGACGCCAAACAGCTCGCTGTCTTGGTGGTCCCGGC 61
QY 21 SerAlaCysGlyLeuGlyPheValProValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 62 TCGGCCTCGCGCTTGGGTTTGTGCGCGTGGTCTACTACAGCCTCTTGTGCTGCTCGGT 121
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAtrArgGlnLys 60
Db 122 TTACAGCAAAATATCTTGGACAGTGTATCTCTCCAGCTGGTGGCAGAGACAGAG 181
QY 61 SerSerTyrAenTyrLeuLeuAlaAlaAlaAspIleLeuValLeuPhePheIle 80
Db 182 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGCCAGACATCTTGGTCTCTTTTTCATA 241
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAenMetGlnMetProGlnValPro 100
Db 242 GTGTTTGGGACTTCTCTTGGAAAGATTTTCATCTTGAACATGACAGATGCCCTCAGGTCCCC 301
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTyrIleThrVal 120
Db 302 GACAGATCATAGAGTGTGGAATTTCTCATCCATCCACACTCCATATGATGATCTACTGTA 361
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 362 CCGTTAACCATTTGACAGGTATATCACTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 421
```



```
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 422 TACCCAGCCGCGACCCGGAAGTCAATTGTAAGTGTTCATCATCACTGCTCTCTGACGACG 481
QY 161 IleProTyrTyrTrpProAlaIleThrThrGluAspTyrIleSerThrSerValHis 180
Db 482 ATCCCCATTACTGGTGGGCCCAACATCTGAGACTGAAGACTTACATCAGCAGCCTCTGTGCAT 541
QY 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 542 CAGTCTCTATCTGGATCACTGCTTACCGTCTACTGTGTCCTGCTCCATCTCTCTTC 601
QY 201 IleLeuAsnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 602 ATCTTGAATCAATCTTGTGTACAAGCTCAGGAGGAGCAATTTCTGCTCTCGTGGC 661
QY 221 TyrSerThrGlyValThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 662 TACTCACGGGGAAGACCCCGCATCTTGTTCACCAATTACCTCCATCTTTGCCACATTT 721
QY 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGCATCATGATCTTTTACCACTTATGGGGCGCCCATCCAGAACCCG 781
QY 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 782 TGGCTGGTACATCATGATCGGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC 841
QY 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 842 AACTTCTCTCTACTGCTTCTATCAGCAAGCGGTTCCGCACCATGCGCAGCCGCGCTC 901
QY 301 LysAlaPhePheLysCysGlnValProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 902 AAGGCTTCTCTAGTGCAGCAAGCACTGTACAGTTCTACCAATCATCAATCTTCTCC 961
QY 321 IleThrSerSerProTrpIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 962 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGTGTAC 1021
QY 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350
Db 1022 CAGTATGACAAAATGGAACCTATAAAGTATCCCGTGATCCCATAGGTGTGGCAAC 1081
QY 350 ----- 350
Db 1082 TACTGCCTCTGTATATCCATTTCCAGATGGGAAGGTGCCATCTATGCTGAGCAGC 1141
QY 350 ----- 350
Db 1142 TCTCTTAAGAGTGTATTCGATTTCTGTCCTCGCAGAGCTGGGCAATCTCAGACTG 1201
QY 350 ----- 350
Db 1202 GTAGATGAGAGAGATGGGAAGAGAAGAGAGAGCATGAAGCTGTTTTTACTTATGCA 1261
QY 351 -----SerArgAsnAspSerLysSerTyrGlnPheGluAspAlaIleG 366
Db 1262 TTTATTTCCACAGATCGTATATGACGAAAGCTCTTACCAGTTTGAAGATGCCATTGG 1321
QY 366 yAlaCysValIleIleLeu 372
Db 1322 AGCTTGTGTCATCATCTCTG 1340
```

RESULT 12

US-10-246-583-21  
; Sequence 21, Application US/10246583  
; Publication NO. US20040058862A1  
; GENERAL INFORMATION:  
; APPLICANT: Majumder  
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME  
; FILE REFERENCE: 15966-729CIP2CON1

; CURRENT APPLICATION NUMBER: US/10/246,583  
; CURRENT FILING DATE: 2002-12-06  
; PRIOR APPLICATION NUMBER: 10/174,364  
; PRIOR FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 128  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 21  
; LENGTH: 1343  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2)..(1060)  
US-10-246-583-21

Alignment Scores:

Pred. No.: 1.41e-184 Length: 1343  
Score: 1903.00 Matches: 371  
Percent Similarity: 83.00% Conservative: 0  
Best Local Similarity: 83.00% Mismatches: 1  
Query Match: 96.75% Indels: 75  
DB: 17 Gaps: 1

US-10-689-832-20 (1-372) x US-10-246-583-21 (1-1343)

```
QY 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGly 20
Db 2 ATGGAGCACACGACGCCACCTCCAGCCAAACAGCTCGTGTCTTGGTGGTCCCCCGC 61
QY 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuGly 40
Db 62 TCGGCTCGGCTGGGTTTCGTGCGCGTGTCTACTACAGCTCTTGTGTGCTCGGT 121
QY 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaArgGlnLys 60
Db 122 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGTGGCAGAGACAGAG 181
QY 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 182 TCTCTCTACACTATCTCTTGGCACTCGTGTGCGCACTCTTGTGCTCTCTTTTTCATA 241
QY 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 242 GTGTTTGGGACTTCTGTTGGAAGATTCATCTTGAACATGCAGATGCCTCAGGTCCC 301
QY 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 302 GACAAGATCATAGAAGTGTGGAAATTCATCTCCATCCACCTCCATATGGATTACTGTA 361
QY 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 362 CCGTTAACCATGTAGAGGTATATCACTGTCTGCGCACCCGCTCAAGTACCACACGGTCTCA 421
QY 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
```

```
Db 422 TACCACGCCGACCCGGAAGTCTATTGTAAGTGTATTACATCACCTGTTCTGACACG 481
Qy 161 IleProTyrTyrTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
Db 482 ATCCCTATTACTGTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 541
Qy 181 HisValLeuIleTrpIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 542 CACGTCCTCATCTGATCCATCGTCTTACCCTGTACCTGGTGGCCGTCTCCATCTCTTC 601
Qy 201 IleLeuAsnSerIleIleValTyrLysLeuArgLysSerAsnPheAsnLeuArgGly 220
Db 602 ATCTTGAACTCAATCTGTGACAGCTCAGAGGAAGCAATTTTCGTCTCGTGGC 661
Qy 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 662 TACTCCACGGGAGACACCGCCATCTTTGTCACCAATTACCTCCATCTTTGCCACACTT 721
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGCATCATCATGATCTTTACCACCTCTATGGGGCGCCCATCCAGAACCGC 781
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 782 TGGCTGGTACACATCATGTCCGACATTCGCAACATGCTAGCCCTTCTGAACACAGCCATC 841
Qy 281 AsnPheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaThrLeu 300
Db 842 AACTTCTTCTCTACTGCTTCATCAGCAAGCGTTTCGCACCATGGCAGCGCCACGGCTC 901
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 902 AAGGCTTTCTTCAAGTGGCAGAACCACTGTACAGTTCACCAATCATCAACTTTTCC 961
Qy 321 IleThrSerSerProTyrIleSerProAlaAsnSerHisCysIleLeuMetLeuValTyr 340
Db 962 ATAACAGTAGCCCTGGATCTCGCGCGCAAACTCACACTGCATCAAGATGCTGTGTATC 1021
Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350
Db 1022 CAGTATGACAAAATGGAAAACCTATAAAAGTATCCCGTGATCCATAGGTGTGGCAAC 1081
Qy 350 ----- 350
Db 1082 TACTGCCTCTGTCTAATCCATTTCCAGATGGGAAGGTGTCCCATCTATGCTGAGCAGC 1141
Qy 350 ----- 350
Db 1142 TCTCCTTAAGAGTGTCTAATCCGATTTCTGTCTCCGCGAGACTGGGCAATTTCTCAGACTG 1201
Qy 350 ----- 350
Db 1202 GTAGATGAGAGAGATGGAAGAGAGAAAGGAGAGCATGAAGCTGTGTTTTTACTTTATGCA 1261
Qy 351 -----SerArgAsnAspSerLysSerSerTyrGlnPheGluAspAlaIleG1 366
Db 1262 TTTATTTCCACAGAGTCGTAATGACAGCAAAAGACTCTACCAGTTTGAAGATGCCAATGG 1321
Qy 366 yAlaCysValIleIleLeu 372
Db 1322 AGCTGTGTGCATCATCCTG 1340
```

## RESULT 13

```
US-10-689-832-21
; Sequence 21, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
```

```
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 21
; LENGTH: 1343
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-689-832-21
Alignment Scores:
Pred. No.: 1,41e-184 Length: 1343
Score: 1903.00 Matches: 371
Percent Similarity: 83.00% Conservative: 0
Best Local Similarity: 83.00% Mismatches: 1
Query Match: 96.75% Indels: 75
DB: 18 Gaps: 1
US-10-689-832-20 (1-372) x US-10-689-832-21 (1-1343)
```

```
Qy 1 MetGluHisThrHisAlaHisLeuAlaAsnSerSerLeuSerTyrTrpSerProGly 20
Db 2 ATGGAGCACAGCAGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGGTCCCCCGC 61
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 62 TCGCCCTCGCGCTTGGGTTTTCGTGCCGTGGTCTACTACAGCCTCTTGTGCTCGGT 121
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAtrArgGlnLys 60
Db 122 TTACAGCAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAG 181
Qy 61 SerSerTyrAsnTyrLeuLeuAlaLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 182 TCCTCTCTACAACTATCTCTTGGCCTCGCTGCTGCCGACATCTTGGTCTCTTTTCATA 241
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleIleuAsnMetGlnMetProGlnValPro 100
Db 242 GTGTTGTGGACTTCCCTGTTGGAAGATTTATCTCTGAACATGACAGATGCCTCAGGTCCCC 301
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleTrpIleThrVal 120
Db 302 GACAAGATCATAGAAGTGTGGAATTTCTATCCATCCACCTCCATATGATGATCTGTA 361
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 362 CCGTTAAACATTGACAGGTATATCACTGTCTGCCACCGCTCAAGTACCAACACGCTCTCA 421
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 422 TACCAGCCCGCACCCGGAAGATCATTTAAGTGTTTACATCATCTGCTGCTCTCTGACCAGC 481
Qy 161 IleProTyrTyrTrpTrpProAsnIleTrpThrGluAspTyrIleSerThrSerValHis 180
```

```
Db 482 ATCCCTATTACTGGTGGCCCAACATCTGCACTGAAGACTACATCAGCACCTCTGTGCAT 541
Qy HisValLeuIleThrPheHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 542 CAGGTCTCTCATCTGGATCACTGCTTACCGTCTACTGCTGCTGCTGCTGCTGCTGCT 601
Qy 201 IleLeuAnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 602 ATCTTGAATCAATCATCTGTACAGCTCAGGAGGAGAGCAATTTCTGCTCCGTGGC 661
Qy 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrIleu 240
Db 662 TACTCCACGGGAGAACCCCGCATCTTGTTCACCATTAACCTTCTTCCACATCTT 721
Qy 241 TrpAlaProArgIleIleMetIleLeuTyrHisLeuTyrGlyAlaProIleGlnAsnArg 260
Db 722 TGGGCCCCCGCATCATCATGATCTTTTACCATCTATGCGGGGCGCCATCCAGAACCG 781
Qy 261 TrpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAsnThrAlaIle 280
Db 782 TGGCTGGTACACATCATGTCGACATTCGCAATGTCACATGCTAGCCCTTCTGAACACAGCCATC 841
Qy 281 AsnPhePheLeuTyrCysPheIleSerLysArgPheArgThrMetAlaAlaAlaThrIleu 300
Db 842 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCCGACCATGCGCAGCCGCAAGTTC 901
Qy 301 LysAlaPhePheLysCysGlnLysGlnProValGlnPheTyrThrAsnHisAsnPheSer 320
Db 902 AGGCTTCTTCAAGTCCAGAGCAACCTGTACAGTTCTACCCATCATCATCTTCTTCTTCC 961
Qy 321 IleThrSerSerProThrIleSerProAlaAsnSerHisCysIleLysMetLeuValTyr 340
Db 962 ATAACAAGTAGCCCTTGGATCTCGCGCGCAAACTCACACTGATCAAGATGCTGGTGATC 1021
Qy 341 GlnTyrAspLysAsnGlyLysProIle-Lys----- 350
Db 1022 CAGTATGACAAAATGGAAAACCTATAAAAGTATCCCCGTGATCCATAGGTGTGGCAAC 1081
Qy 350 ----- 350
Db 1082 TACTGCCTCTGCTAATCACTTCAGATGGGAAGGTGCCATCTATGGCTGAGCAGC 1141
Qy 350 ----- 350
Db 1142 TCTCCTTAAGAGTGCTAATCCGATTCCTGCTCTCCGAGACTGGGCAATCTCAGACTG 1201
Qy 350 ----- 350
Db 1202 GTAGATGAGAGAGATGGGAAGAGAAAGAGAGCATGAAGCTTTGTTTTTATTATGCA 1261
Qy 351 -----SerArgAsnAspSerLysSerTyrGlnPheGluAspAlaIleGly 366
Db 1262 TTTATTTCCAGAGTCTGTAATGACAGCAAAAGCTCTTACCAGTTTGAAGATGCCATTGG 1321
Qy 366 yAlaCysValIleIleLeu 372
Db 1322 AGCTGTGTGCATCATCTG 1340

RESULT 14
US-10-012-140-10
; Sequence 10, Application US/10012140
; Publication No. US20030009017A1
; GENERAL INFORMATION:
; APPLICANT: Leiby, Kevin R.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 38650, 28472, 5495, 65507, 81588, AND
; TITLE OF INVENTION: 14354 METHODS AND COMPOSITIONS OF HUMAN PROTEINS AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 381552004900
; CURRENT APPLICATION NUMBER: US/10/012,140
; CURRENT FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: 60/246,768
```

```
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,772
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,185
; PRIOR FILING DATE: 2000-11-15
; NUMBER OF SEQ ID NOS: 49
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 1526
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (139)...(1200)
US-10-012-140-10

Alignment Scores:
Pred. No.: 4,97e-182 Length: 1526
Score: 1879.00 Matches: 366
Percent Similarity: 82.33% Conservativity: 2
Best Local Similarity: 81.88% Mismatches: 4
Query Match: 95.53% Indels: 75
DB: 14 Gaps: 1

US-10-689-832-20 (1-372) x US-10-012-140-10 (1-1526)
Qy 1 MetGluHisThrHisAlaHisLeuAlaAlaAsnSerSerLeuSerTyrTrpSerProGly 20
Db 139 ATGGAGCAACAGCAGCCACCCTCGCAGCAACAGCTCGCTGCTTGTGGTGGTCCCCGGC 198
Qy 21 SerAlaCysGlyLeuGlyPheValProValValTyrTyrSerLeuLeuLeuCysLeuGly 40
Db 199 TCGGCTCGCGCTTGGGTTTGTGTCCTGCTGCTACTACAGCCTCTTGCAGTGCCTCGT 258
Qy 41 LeuProAlaAsnIleLeuThrValIleIleLeuSerGlnLeuValAlaAlaArgGlnLys 60
Db 259 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGCTGGCAGAGACAGAAG 318
Qy 61 SerSerTyrAsnTyrLeuLeuAlaAlaAspIleLeuValLeuPhePheIle 80
Db 319 TCCTCTCAAACTATCTCTTGGCACTCGCTGCTGCGACATCTTGGTCTCTTTTTCATA 378
Qy 81 ValPheValAspPheLeuLeuGluAspPheIleLeuAsnMetGlnMetProGlnValPro 100
Db 379 GTGTTTGGAGCTTCTCTGTGGAAGATTTTCATTTTGAACATGCAGATGCCTCAGTCCC 438
Qy 101 AspLysIleIleGluValLeuGluPheSerSerIleHisThrSerIleThrIleThrVal 120
Db 439 GACAGATCATAGAGTCTGGAATTTCTCATCCATCCACCTCCATATGGATTTACTGTA 498
Qy 121 ProLeuThrIleAspArgTyrIleAlaValCysHisProLeuLysTyrHisThrValSer 140
Db 499 CCGTTAACCATTCAGAGGTATATCGCTGTGCGCACCCGCTCAAGTACCACAGGCTCTCA 558
Qy 141 TyrProAlaArgThrArgLysValIleValSerValTyrIleThrCysPheLeuThrSer 160
Db 559 TACCAGCCCGCACCCGGAAGTCAATGTAAGTGTATACATCACCTGCTTCTGTGACACG 618
Qy 161 IleProTyrTyrTrpTrpProAsnIleThrGluAspTyrIleSerThrSerValHis 180
Db 619 ATCCCTATTACTGTTGGCCCAACATCTCGAGCTGAAGACTACATCAGCACCTCTGTGCAT 678
Qy 181 HisValLeuIleThrIleHisCysPheThrValTyrLeuValProCysSerIlePhePhe 200
Db 679 CACGTCCTCATCTGGGTCCACTGCTTACCCTGCTTACCAGGTGCCCTGCTCCATCTTCTTC 738
Qy 201 IleLeuAnSerIleIleValTyrLysLeuArgArgLysSerAsnPheArgLeuArgGly 220
Db 739 ATCTTGAATCAATCGTTGTGTACAGCTCAGGAGGAGAGCAATTTCTGCTCTGTGGC 798
Qy 221 TyrSerThrGlyLysThrAlaIleLeuPheThrIleThrSerIlePheAlaThrLeu 240
Db 799 TACTCCAGGGGAGAGACACCGCCATCTTGTTCACCATTAACCTCCATCTTTGCCATCTT 858
```

Qy	241	TtpAlaproArgllelMetilleLeutyHisleutyrglyAlaProIleGlnAenArg	260
Db	859	TGGGCCCCCGCATCATCATGATTCTTTACCACTCTATGGGCGCCATCCAGAACCGC	918
Qy	261	TtpLeuValHisIleMetSerAspIleAlaAsnMetLeuAlaLeuLeuAenThrAlaIle	280
Db	919	TGGCTGGTACATCATGTCCGACATTCGCACATGCTAGCCCTTCGAAACAGACGATC	978
Qy	281	AsnPhePheLeuTyrcysPheIleSerLysArgPheArghrMetAlaAlaThrLeu	300
Db	979	AAC TTCCTCTACTCTTCATCAGCAAGCGGTTCGCACCATTGGCAGCGCACGCTC	1038
Qy	301	LysAlaPhePheLysCysGlnLysGlnProValcInPheTyrrThrAsnHisAenPheSer	320
Db	1039	ARGCTTCTTCAAGTGCCAGAACAACCTGTACAGTTCTACACCAATCATAACTTTTCC	1098
Qy	321	IleThrSerSerProIleSerProAlaAsnSerHisCysIleLysMetLeuValTyrr	340
Db	1099	ATAACAGTAGCCCCCTGGATCTCCCGGCAAACTCACACTGCATCAAGATGCTGGTGATC	1158
Qy	341	GlnTyrrAspLysAsnGlyLysProIle-Lys-----	350
Db	1159	CAGTATGACAAAATGGAACACCTATAAAAGTATCCCCTGATTCCTAGGTGGCAAC	1218
Qy	350	-----	350
Db	1219	TACTGCCTCTGTCTTAATCCATTTCCAGATGGGAAGGTGCCCATCTATGGCTGAGCAGC	1278
Qy	350	-----	350
Db	1279	TCTCCTTAAGATGCTAATCGGATTTCTCTCTCCCGCAGACTGGGCAATTCACAGACTG	1338
Qy	350	-----	350
Db	1339	GTAGATGAGAAGAGATGGAAGAGAGAAAAAGAGAGCATGAAGCTGTTTTTACTTATGCA	1398
Qy	351	-----SerArgenAspSeryLysSerSerTyrrGlnPhecluAspAlarIeGI	366
Db	1399	TTTTATTCCACAGATCGTAATGACAAAGCTCTACAGTTTGNAGATGCCATTGG	1458
Qy	366	yAlaCysValIleIleLeu 372	
Db	1459	AGCTTGTGTCATCATCTCTG 1477	

## RESULT 15

```

US-10-450-590-1
; Sequence 1, Application US/10450590
; Publication No. US20040076985A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: REGULATION OF HUMAN CHEMOKINE-LIKE RECEPTOR
; FILE REFERENCE: LI0316 Foreign Countries
; CURRENT APPLICATION NUMBER: US/10/450,590
; CURRENT FILING DATE: 2003-06-16
; PRIOR APPLICATION NUMBER: US 60/254,923
; PRIOR FILING DATE: 2000-12-14
; PRIOR APPLICATION NUMBER: US 60/280,110
; PRIOR FILING DATE: 2001-04-02
; PRIOR APPLICATION NUMBER: US 60/299,474
; PRIOR FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1059
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-450-590-1

```

Alignment Scores:	
Pred. No.:	5.08e-180
Score	1857.00
Percent Similarity:	100.00%
Conservative:	0
Length:	1059
Matches:	350

Qy 341 GlnTyrAspLysAsnGlyLysProIleLys 350  
|||  
Db 1021 CAGTATGACAAAAATGAAAACTATATAA 1050

Search completed: February 12, 2005, 16:03:35  
Job time : 608 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 12, 2005, 10:40:56 ; Search time 659 Seconds  
(without alignments)  
10019.142 Million cell updates/sec

Title: US-10-689-832-19  
Perfect score: 1119  
Sequence: 1 atggagcacacgcacgccca.....cttgtgtcatcatctgtga 1119

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 5378673 seqs, 2950229984 residues

Total number of hits satisfying chosen parameters: 10757346

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.\*  
1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq.\*  
16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq.\*  
17: /cgn2\_6/ptodata/1/pubpna/US10E\_PUBCOMB.seq.\*  
18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq.\*  
19: /cgn2\_6/ptodata/1/pubpna/US10\_NEW\_PUB.seq.\*  
20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq.\*  
21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq.\*  
22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1119	100.0	1119	10	US-09-813-432-19
2	1119	100.0	1119	17	US-10-174-364-19
3	1119	100.0	1119	17	US-10-246-583-19
4	1119	100.0	1119	18	US-10-689-832-19
5	1119	100.0	1130	17	US-10-333-946-25
6	1119	100.0	2198	16	US-10-314-076-1
7	1091	97.5	1110	16	US-10-314-076-3
8	1091	97.5	2189	14	US-10-219-834-6
9	1050	93.8	1059	17	US-10-450-590-1
10	1050	93.8	1062	17	US-10-296-294A-3
11	1050	93.8	1202	14	US-10-094-417-1
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 19, Appl
					Sequence 25, Appl
					Sequence 1, Appl
					Sequence 3, Appl
					Sequence 6, Appl
					Sequence 1, Appl
					Sequence 3, Appl
					Sequence 1071, Ap
					Sequence 909, App
					Sequence 62, Appl

12	1048.4	93.7	1062	9	US-09-995-225-15	Sequence 15, Appl
13	1048.4	93.7	1062	10	US-09-995-225-15	Sequence 15, Appl
14	1048.4	93.7	1062	17	US-10-296-294A-4	Sequence 4, Appl
15	1048.4	93.7	1343	10	US-09-813-432-21	Sequence 21, Appl
16	1048.4	93.7	1343	17	US-10-174-364-21	Sequence 21, Appl
17	1048.4	93.7	1343	17	US-10-246-583-21	Sequence 21, Appl
18	1048.4	93.7	1343	17	US-10-689-832-21	Sequence 21, Appl
19	1048.4	93.7	2117	18	US-10-779-104-1	Sequence 1, Appl
20	1040.4	93.0	1062	14	US-10-012-140-12	Sequence 12, Appl
21	1040.4	93.0	1526	14	US-10-012-140-10	Sequence 10, Appl
22	934.8	83.5	1062	17	US-10-450-590-9	Sequence 9, Appl
23	929.8	83.1	957	17	US-10-174-364-84	Sequence 84, Appl
24	929.8	83.1	957	17	US-10-246-583-84	Sequence 84, Appl
25	926	82.8	1002	15	US-10-343-650A-21	Sequence 25, Appl
26	926	82.8	1158	17	US-10-079-384-25	Sequence 21, Appl
27	922.4	82.4	1032	17	US-10-450-590-5	Sequence 5, Appl
28	922.4	82.4	1070	17	US-10-450-590-4	Sequence 4, Appl
29	922.4	82.4	1826	17	US-10-450-590-6	Sequence 6, Appl
30	844.2	75.4	864	10	US-09-791-932-6	Sequence 6, Appl
31	832	74.4	1038	14	US-10-094-417-19	Sequence 19, Appl
32	756.4	67.6	795	15	US-10-225-567A-586	Sequence 586, App
33	524	46.8	930	10	US-09-791-932-3	Sequence 4, Appl
34	268.2	24.0	1466	17	US-10-328-916-4	Sequence 4, Appl
35	268.2	24.0	1962	17	US-10-333-946-29	Sequence 29, Appl
36	266.6	23.8	1125	9	US-09-995-225-5	Sequence 5, Appl
37	266.6	23.8	1125	10	US-09-995-225-5	Sequence 15, Appl
38	266.6	23.8	1125	14	US-10-012-140-15	Sequence 13, Appl
39	266.6	23.8	1427	14	US-10-094-417-13	Sequence 1, Appl
40	266.6	23.8	1530	13	US-10-011-147-1	Sequence 1, Appl
41	266.6	23.8	1719	14	US-10-012-140-13	Sequence 13, Appl
42	266.6	23.8	1793	17	US-10-328-916-3	Sequence 3, Appl
43	243.8	21.8	34118	15	US-10-017-161-1071	Sequence 1071, Ap
44	243.8	21.8	34118	17	US-10-292-798-909	Sequence 909, App
45	80	7.1	80	14	US-10-219-834-62	Sequence 62, Appl

ALIGNMENTS

RESULT 1

US-09-813-432-19  
; Sequence 19, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majmuder, Kamud  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1 Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,664  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,836

```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/193,843
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 78
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-813-432-19

Query Match      100.0%; Score 1119; DB 10; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACAGCAGCCACCTCGAGCCAAACAGCTGCTGCTCTTGGTGGTCCCGGC 60
Db 1 ATGGAGCACAGCAGCCACCTCGAGCCAAACAGCTGCTGCTCTTGGTGGTCCCGGC 60

QY 61 TCGGCTCGGCTTGGGTTTCGTCGCCGCGTCTACTACAGCTCTTGTGTCGCTCGGT 120
Db 61 TCGGCTCGGCTTGGGTTTCGTCGCCGCGTCTACTACAGCTCTTGTGTCGCTCGGT 120

QY 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
Db 121 TTACCAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180

QY 181 TCCTCTCAAACTATCTCTTGGGACATCGTGTGCGGACATCTTGGTCTCTTTTTCATA 240
Db 181 TCCTCTCAAACTATCTCTTGGGACATCGTGTGCGGACATCTTGGTCTCTTTTTCATA 240

QY 241 GGTGTTGGGACTCTCTGTTGAAGATTTTCATCTTGACATGCAGATGCTCAGTCCCC 300
Db 241 GGTGTTGGGACTCTCTGTTGAAGATTTTCATCTTGACATGCAGATGCTCAGTCCCC 300

QY 301 GACAAAGATCATAGAAAGTCTGCGAAATTTCTCATCCATCCACCTCCATATGAACTGTA 360
Db 301 GACAAAGATCATAGAAAGTCTGCGAAATTTCTCATCCATCCACCTCCATATGAACTGTA 360

QY 361 CGGTTAACATTTGACAGGTATATCGTGTCTGCGCACCGCTCAAGTACACAGGTCTCA 420
Db 361 CGGTTAACATTTGACAGGTATATCGTGTCTGCGCACCGCTCAAGTACACAGGTCTCA 420

QY 421 TACCCAGCCCGCACCGGAAAGTCTATGTAAGTGTATCATCACTGCTTCTGACCGAGC 480
Db 421 TACCCAGCCCGCACCGGAAAGTCTATGTAAGTGTATCATCACTGCTTCTGACCGAGC 480

QY 481 ATCCCTCTATTACTGGTGGCCCAACATCTGGAAGTGAAGTACATCAGCACCTCTGTGCAT 540
Db 481 ATCCCTCTATTACTGGTGGCCCAACATCTGGAAGTGAAGTACATCAGCACCTCTGTGCAT 540

QY 541 CAGCTGCTCATCTGGATCCACTGCTTACCGTCTACCTGCTGCTGCTGCTGCTGCTTTC 600
Db 541 CAGCTGCTCATCTGGATCCACTGCTTACCGTCTACCTGCTGCTGCTGCTGCTTTC 600

QY 601 ATCTTGAATCAATCAATGTTGACAGCTCAGGAGAGAGCAATTTTCGTCCTCGTGGC 660
Db 601 ATCTTGAATCAATCAATGTTGACAGCTCAGGAGAGAGCAATTTTCGTCCTCGTGGC 660

QY 661 TACTCCAGGGGAAGACCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720
Db 661 TACTCCAGGGGAAGACCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720

QY 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGCGGCGCCCATCCAGAACCGC 780
Db 721 TGGGCCCCCGCATCATCATGATTTTACCACTCTATGCGGCGCCCATCCAGAACCGC 780

QY 781 TGGCTGGTACATCATATGTCGACATTCGACATGCTAGCCCTTCTGACACAGCCATC 840
Db 781 TGGCTGGTACATCATATGTCGACATTCGACATGCTAGCCCTTCTGACACAGCCATC 840

QY 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGCGACCATGGCAGCCCGCAGCTC 900
Db 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGCGACCATGGCAGCCCGCAGCTC 900
```

```
Db 841 AACTTCTTCTCTACTGCTTCTATCAGCAAGCGGTTCGCGACCATGGCAGCCCGCAGCTC 900
QY 901 AAGGCTTTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960
Db 901 AAGGCTTTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960
QY 961 ATAACAAGTAGCCCTCGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGCTGTAC 1020
Db 961 ATAACAAGTAGCCCTCGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGCTGTAC 1020
QY 1021 CAGTATGACAAAATGGAAAACCTATAAAAAAGTCTGTAATGACAGCAAAAGCTCTTACCAG 1080
Db 1021 CAGTATGACAAAATGGAAAACCTATAAAAAAGTCTGTAATGACAGCAAAAGCTCTTACCAG 1080
QY 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGCA 1119
Db 1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGCA 1119
```

RESULT 2

```
US-10-174-364-19
; Sequence 19, Application US/10174364
; Publication No. US20030216308A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2
; CURRENT APPLICATION NUMBER: US/10/174,364
; CURRENT FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,836
; PRIOR FILING DATE: 2000-03-29
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
US-10-174-364-19
```

```
Query Match      100.0%; Score 1119; DB 17; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGCACAGCAGCCACCTCGAGCCAAACAGCTGCTGCTCTTGGTGGTCCCGGC 60
Db 1 ATGAGCACAGCAGCCACCTCGAGCCAAACAGCTGCTGCTCTTGGTGGTCCCGGC 60

QY 61 TCGGCTCGGCTTGGGTTTCGTCGCCGCGTCTACTACAGCTCTTGTGTCGCTCGGT 120
Db 61 TCGGCTCGGCTTGGGTTTCGTCGCCGCGTCTACTACAGCTCTTGTGTCGCTCGGT 120
```



```

QY 121 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
Db 121 TTACGAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
QY 181 TCCTCTCAACAATCTCTTGACACTCGCTGCTGCGACATCTTGGTCTCTTTTTCATA 240
Db 181 TCCTCTCAACAATCTCTTGACACTCGCTGCTGCGACATCTTGGTCTCTTTTTCATA 240
QY 241 GTGTTGTGGACTTCTGTTGGAAGATTTCACTTTTGAACATGACAGTGCCTCAGGTCCC 300
Db 241 GTGTTGTGGACTTCTGTTGGAAGATTTCACTTTTGAACATGACAGTGCCTCAGGTCCC 300
QY 301 GACAGATCATAGAGTGTGAAATCTCATCTCCATCCACCTCCATATGATGATCTGTA 360
Db 301 GACAGATCATAGAGTGTGAAATCTCATCTCCATCCACCTCCATATGATGATCTGTA 360
QY 361 CGTTAAACCAATGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGCTCTCA 420
Db 361 CGTTAAACCAATGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGCTCTCA 420
QY 421 TACCAGCCCGCACCGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480
Db 421 TACCAGCCCGCACCGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480
QY 481 ATCCCTATTAATGCTGGCCCAACATCTGGAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 540
Db 481 ATCCCTATTAATGCTGGCCCAACATCTGGAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 540
QY 541 CAGGCTCTATCTGATGATGCTTCAACGCTACCTGCTGGTGGCCCTGCTTCTTCTTC 600
Db 541 CAGGCTCTATCTGATGATGCTTCAACGCTACCTGCTGGTGGCCCTGCTTCTTCTTC 600
QY 601 ATCTTGAACATCAATGTTGTACAGCTCAGGAGGAGCAATTTTCTGCTCCGCTGGC 660
Db 601 ATCTTGAACATCAATGTTGTACAGCTCAGGAGGAGCAATTTTCTGCTCCGCTGGC 660
QY 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACATTAATCTCAATCTTTGGCCACAT 720
Db 661 TACTCCAGGGGAGACACCGCCATCTTGTTCACATTAATCTCAATCTTTGGCCACAT 720
QY 721 TGGGCCCCCGCATCATATGATTTCTTACACCTCTATGGGCGCCCATCCAGAACCG 780
Db 721 TGGGCCCCCGCATCATATGATTTCTTACACCTCTATGGGCGCCCATCCAGAACCG 780
QY 781 TGGCTGTGTACATCATGTTCCGACATTTGCAACATGCTAGCCCTTCTGAAACACAGCCATC 840
Db 781 TGGCTGTGTACATCATGTTCCGACATTTGCAACATGCTAGCCCTTCTGAAACACAGCCATC 840
QY 841 AACTTCTTCTCTACTGCTTATAGCAAGCGGTTCCGACCATGGCAGCGCCACAGCTC 900
Db 841 AACTTCTTCTCTACTGCTTATAGCAAGCGGTTCCGACCATGGCAGCGCCACAGCTC 900
QY 901 AAGGCTTTCTCAAGTGCAGAGCAACTGTGACAGTTCTACAGTTCTACCAATCAATCTTCC 960
Db 901 AAGGCTTTCTCAAGTGCAGAGCAACTGTGACAGTTCTACAGTTCTACCAATCAATCTTCC 960
QY 961 ATAACAAGTAGCCCTGATCTGCGCGCAACTCACAAGTCAAGTGTGCTGTAC 1020
Db 961 ATAACAAGTAGCCCTGATCTGCGCGCAACTCACAAGTCAAGTGTGCTGTAC 1020
QY 1021 CAGTATGACAAAATGGAACCTATAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 1080
Db 1021 CAGTATGACAAAATGGAACCTATAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 1080
QY 1081 TTTGAAGTGCATTTGGAGCTTGTGTCATCTCTGTA 1119
Db 1081 TTTGAAGTGCATTTGGAGCTTGTGTCATCTCTGTA 1119

```

RESULT 3

US-10-246-583-19

; Sequence 19, Application US/10246583

; Publication No. US20040058862A1

```

; GENERAL INFORMATION:
; APPLICANT: Majumder
; TITLE OF INVENTION: NOVEL POLYPEPTIDES AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 15966-729CIP2CON1
; CURRENT APPLICATION NUMBER: US/10/246,583
; CURRENT FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: 10/174,364
; PRIOR FILING DATE: 2002-06-17
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,664
; PRIOR FILING DATE: 2000-03-28
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 1119
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1116)
; US-10-246-583-19

```

```

Query Match 100.0%; Score 1119; DB 17; Length 1119;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGGACACAGCAGCCACCTCGAGCCAAAGCTCGCTGCTTGTGGTGTCCCGGCG 60
Db 1 ATGAGGACACAGCAGCCACCTCGAGCCAAAGCTCGCTGCTTGTGGTGTCCCGGCG 60
QY 61 TCGGCTCGGCTTGGGTTTGTGCGCGTGTCTACTACGCTCTTGTGCTGCTCGGT 120
Db 61 TCGGCTCGGCTTGGGTTTGTGCGCGTGTCTACTACGCTCTTGTGCTGCTCGGT 120
QY 121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
Db 121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180
QY 181 TCCTCTCAACAATCTCTTGGCACTGCTGCTGCGCAATCTTGGTCTCTTTTTCATA 240
Db 181 TCCTCTCAACAATCTCTTGGCACTGCTGCTGCGCAATCTTGGTCTCTTTTTCATA 240
QY 241 GTGTTGTGGACTTCTGTTGGAAGATTTCACTTTTGAACATGACAGTGCCTCAGGTCCC 300
Db 241 GTGTTGTGGACTTCTGTTGGAAGATTTCACTTTTGAACATGACAGTGCCTCAGGTCCC 300
QY 301 GACAGATCATAGAGTGTGAAATCTCATCTCCATCCACCTCCATATGATGATCTGTA 360
Db 301 GACAGATCATAGAGTGTGAAATCTCATCTCCATCCACCTCCATATGATGATCTGTA 360
QY 361 CCGTTAAACCAATGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGCTCTCA 420
Db 361 CCGTTAAACCAATGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGCTCTCA 420
QY 421 TACCAGCCCGCACCGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480
Db 421 TACCAGCCCGCACCGGAAAGTCAATGTAAGTGTTCATCACCTGCTTCTGACACG 480

```

Qy	481	ATCCCTATTACTGGTGGCCCAACATCTGGA	CTGAAGCTACATCAGGACCTCTGTGCAT	540
Db	481	ATCCCTATTACTGGTGGCCCAACATCTGGA	CTGAAGCTACATCAGCACCCTCTGTGCAT	540
Qy	541	CAGTCCCTCATCTGGATCCACTGCTTCA	CCGCTCTACCTGGTGGCCCTGCTCCATCTTCTTC	600
Db	541	CAGTCCCTCATCTGGATCCACTGCTTCA	CCGCTCTACCTGGTGGCCCTGCTCCATCTTCTTC	600
Qy	601	ATCTTGAATCAATCATTTGTGTACAAGCT	CAGGAGGAAGCAATTTTCTGCTCCGTGGC	660
Db	601	ATCTTGAATCAATCATTTGTGTACAAGCT	CAGGAGGAAGCAATTTTCTGCTCCGTGGC	660
Qy	661	TACTCCACGGGGAAGACACCGCCATCT	TGTTACCACTTACCTCCATCTTTTGCCACACTT	720
Db	661	TACTCCACGGGGAAGACACCGCCATCT	TGTTACCACTTACCTCCATCTTTTGCCACACTT	720
Qy	721	TGGGCCCCCGCATCATGATCTTTTACA	CTCTATGGGGGCCCATCCAGAACCGC	780
Db	721	TGGGCCCCCGCATCATGATCTTTTACA	CTCTATGGGGGCCCATCCAGAACCGC	780
Qy	781	TGGCTGTGTACACATCATGTCCGACAT	TGCCAATGCTAGCCCTTCTGAAACAGACCATC	840
Db	781	TGGCTGTGTACACATCATGTCCGACAT	TGCCAATGCTAGCCCTTCTGAAACAGACCATC	840
Qy	841	AACTTCTTCCTCTACTGCTTCA	TGCAAGCGGTTCCGACCATGGCAGCGCCACGCTC	900
Db	841	AACTTCTTCCTCTACTGCTTCA	TGCAAGCGGTTCCGACCATGGCAGCGCCACGCTC	900
Qy	901	AAGCTTTTCTTCAAGTCCGAGAACCT	GTACAGTTCTACACCAATCATACTTTTCC	960
Db	901	AAGCTTTTCTTCAAGTCCGAGAACCT	GTACAGTTCTACACCAATCATACTTTTCC	960
Qy	961	ATAACAGTAGCCCCCTGGATCTCG	CCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC	1020
Db	961	ATAACAGTAGCCCCCTGGATCTCG	CCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC	1020
Qy	1021	CAGTATGACAAAATGAAAACCTTATA	AAAGTGTAAATGACGAAGAAGTCTCTACAG	1080
Db	1021	CAGTATGACAAAATGAAAACCTTATA	AAAGTGTAAATGACGAAGAAGTCTCTACAG	1080
Qy	1081	TTTGAAGATGCCATTGGAGCTTGTG	TATCATCTCTGTGA	1119
Db	1081	TTTGAAGATGCCATTGGAGCTTGTG	TATCATCTCTGTGA	1119

## RESULT 4

```

US-10-689-832-19
; Sequence 19, Application US/10689832
; Publication No. US20040121380A1
; GENERAL INFORMATION:
; APPLICANT: Majmuder, Kamud
; TITLE OF INVENTION: Novel Polypeptides and Amino Acids Encoding Same
; FILE REFERENCE: 15966-729DIV1
; CURRENT APPLICATION NUMBER: US/10/689,832
; CURRENT FILING DATE: 2003-10-20
; PRIOR APPLICATION NUMBER: 09/813,432
; PRIOR FILING DATE: 2001-03-20
; PRIOR APPLICATION NUMBER: 60/190,835
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,768
; PRIOR FILING DATE: 2000-03-20
; PRIOR APPLICATION NUMBER: 60/190,972
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,199
; PRIOR FILING DATE: 2000-03-22
; PRIOR APPLICATION NUMBER: 60/191,947
; PRIOR FILING DATE: 2000-03-24
; PRIOR APPLICATION NUMBER: 60/192,665
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,657
; PRIOR FILING DATE: 2000-03-28
; PRIOR APPLICATION NUMBER: 60/192,984

```

Db 841 AACTCTTCTCTTACTGCTTTCATCAGCAAGCGGTTCCGACACATGGCAGCGCCACGCTC 900  
Qy 901 AAGGCTTTCTTCAAGTGCCAGAGCAACCTGTACAGTTCTTACACCAATCATAACTTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGCCAGAGCAACCTGTACAGTTCTTACACCAATCATAACTTTTCC 960  
Qy 961 ATAAAGTAGAGCCCTGGATCTCGCCGGCAAACTCACTGATCAAGATGCTGGTGTAC 1020  
Db 961 ATAAAGTAGAGCCCTGGATCTCGCCGGCAAACTCACTGATCAAGATGCTGGTGTAC 1020  
Qy 1021 CAGTATGACAAAATGGAACCTATAAAAGTCGTAATGACAGCAAAAAGCTCCCTACCAG 1080  
Db 1021 CAGTATGACAAAATGGAACCTATAAAAGTCGTAATGACAGCAAAAAGCTCCCTACCAG 1080  
Qy 1081 TTTGAAGATGCCATTGGAGCTTGTGTCACTCCTGTGA 1119  
Db 1081 TTTGAAGATGCCATTGGAGCTTGTGTCACTCCTGTGA 1119

RESULT 5  
US-10-333-946-25  
; Sequence 25, Application US/10333946  
; Publication No. US20040023252A1  
; GENERAL INFORMATION:  
; APPLICANT: INCYTE GENOMICS, INC.; THORNTON, Michael B.  
; APPLICANT: ARVIZU, Chandra S.; IAL, Preeti G.  
; APPLICANT: BURFORD, Neil; YUE, Henry  
; APPLICANT: GANDHI, Ameena R.; ELLIOTT, Vicki S.  
; APPLICANT: RAMKUMAR, Javalaxmi; BAUGHN, Mariah R.  
; APPLICANT: KALLICK, Deborah A.; CHAWLA, Narinder K.  
; APPLICANT: HAFALIA, April J.A.; YAO, Monique G.  
; APPLICANT: LU, Yan; TRIBOULEY, Catherine M.  
; APPLICANT: POLICKY, Jennifer L.; KEARNEY, Liam  
; APPLICANT: GRAUL, Richard C.; WARREN, Bridget A.  
; APPLICANT: LEE, Ernestine A.; DING, Li  
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTORS  
; FILE REFERENCE: PI-0176 USN  
; CURRENT APPLICATION NUMBER: US/10/333,946  
; CURRENT FILING DATE: 2003-01-22  
; PRIOR APPLICATION NUMBER: PCT/US01/23433  
; PRIOR FILING DATE: 2001-07-25  
; PRIOR APPLICATION NUMBER: US 60/221,478  
; PRIOR FILING DATE: 2000-07-27  
; PRIOR APPLICATION NUMBER: US 60/223,268  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: US 60/227,054  
; PRIOR FILING DATE: 2000-08-21  
; PRIOR APPLICATION NUMBER: US 60/231,121  
; PRIOR FILING DATE: 2000-09-08  
; PRIOR APPLICATION NUMBER: US 60/232,243  
; PRIOR FILING DATE: 2000-09-13  
; PRIOR APPLICATION NUMBER: US 60/232,691  
; PRIOR FILING DATE: 2000-09-15  
; PRIOR APPLICATION NUMBER: US 60/235,146  
; PRIOR FILING DATE: 2000-09-22  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PERL Program  
; SEQ ID NO 25  
; LENGTH: 1130  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No. US20040023252A1 7474767CB1  
US-10-333-946-25

Query Match 100.0%; Score 1119; DB 17; Length 1130;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 ATGGAGCACGACGCCCACTCGCAGGCAACAGCTCTGTGGTGGTCCCCCGGC 60  
|||||

Db 12 ATGGAGCACGACGCCCACTCGCAGGCAACAGCTCTGTGGTGGTCCCCCGGC 71  
Qy 61 TCGGCTCGCGCTTGGGTTTCGTGCCCGGTGTACTACTCAGAGCTCTTGTGTGGTCCGCTCGGT 120  
Db 72 TCGGCTCGCGCTTGGGTTTCGTGCCCGGTGTACTACTCAGAGCTCTTGTGTGGTCCGCTCGGT 131  
Qy 121 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGAAG 180  
Db 132 TTACAGCAAAATATCTTGACAGTGATCATCTCTCCAGCTGGTGGCAAGAGACAGAAG 191  
Qy 181 TCCTCTTCAAACTATCTCTTGGCACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240  
Db 192 TCCTCTTCAAACTATCTCTTGGCACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 251  
Qy 241 GTGTTGTGGACTCTCTGTGGAAGATTTCACTTTGAAACATGACAGATGCTCAGGTCCCC 300  
Db 252 GTGTTGTGGACTCTCTGTGGAAGATTTCACTTTGAAACATGACAGATGCTCAGGTCCCC 311  
Qy 301 GACAAGATCATAGAAGTGTGGAATTTCTCATCTCCATCCACACCTCCATATGGATTACTGTA 360  
Db 312 GACAAGATCATAGAAGTGTGGAATTTCTCATCTCCATCCACACCTCCATATGGATTACTGTA 371  
Qy 361 CGGTTAAACCATTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 420  
Db 372 CGGTTAAACCATTGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACACGGTCTCA 431  
Qy 421 TACCAGCCCGCACCCGGAAGTCTATGTAAGTGTGTTATCATCACCTGTCTCTGACACAGC 480  
Db 432 TACCAGCCCGCACCCGGAAGTCTATGTAAGTGTGTTATCATCACCTGTCTCTGACACAGC 491  
Qy 481 ATCCCTTATTACTGTGGCCCAACATCTGGACTGAAGACTACATCAGACACCTCTGTGCAT 540  
Db 492 ATCCCTTATTACTGTGGCCCAACATCTGGACTGAAGACTACATCAGACACCTCTGTGCAT 551  
Qy 541 CAGGCTCTCATCTGATATCCATCTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600  
Db 552 CAGGCTCTCATCTGATCACTGCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 611  
Qy 601 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTCTGCTCCGTGGC 660  
Db 612 ATCTTGAACCTCAATCATTTGTGTACAGCTCAGAGGAGAGCAATTTTCTGCTCCGTGGC 671  
Qy 661 TACTCCAGGGGAGAACACACCGCCATCTTGTTCACATTTACCTCATCTTTGGCCACACTT 720  
Db 672 TACTCCAGGGGAGAACACACCGCCATCTTGTTCACATTTACCTCATCTTTGGCCACACTT 731  
Qy 721 TGGGCCCCCGGCATCATCATGATTTCTTACCACTCTATGGGGCGCCCATCCAGAACCGC 780  
Db 732 TGGGCCCCCGGCATCATCATGATTTCTTACCACTCTATGGGGCGCCCATCCAGAACCGC 791  
Qy 781 TGGCTGGTACACATCATGTCCGACATTTGCCAATGCGCTAGCCCTTTCTGAACACAGCCATC 840  
Db 792 TGGCTGGTACACATCATGTCCGACATTTGCCAATGCGCTAGCCCTTTCTGAACACAGCCATC 851  
Qy 841 AACTTCTTCTTACTGCTTCAAGCAAGCGGTTCCGACCATGGCAGCGCCACAGCTC 900  
Db 852 AACTTCTTCTTACTGCTTCAAGCAAGCGGTTCCGACCATGGCAGCGCCACAGCTC 911  
Qy 901 AAGGCTTTCTTCAAGTGCCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960  
Db 912 AAGGCTTTCTTCAAGTGCCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 971  
Qy 961 ATAAAGTAGAGCCCTGGATCTCGCCGGCAAACTCACACTGATCAAGATGCTGGTGTAC 1020  
Db 972 ATAAAGTAGAGCCCTGGATCTCGCCGGCAAACTCACACTGATCAAGATGCTGGTGTAC 1031  
Qy 1021 CAGTATGACAAAATGGAACCTATAAAAGTCTGTAATGACAGCAAAAAGCTCTTACCAG 1080  
Db 1032 CAGTATGACAAAATGGAACCTATAAAAGTCTGTAATGACAGCAAAAAGCTCTTACCAG 1091  
Qy 1081 TTTGAAGATGCCATTGGAGCTTGTGTCACTCCTGTGA 1119  
Db 1092 TTTGAAGATGCCATTGGAGCTTGTGTCACTCCTGTGA 1130

```
RESULT 6
US-10-314-076-1
; Sequence 1, Application US/10314076
; Publication No. US20030152977A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRMY34, AND VARIANTS A
; FILE REFERENCE: D0197NP
; CURRENT APPLICATION NUMBER: US/10/314,076
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: U.S. 60/338,371
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 2198
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1047)..(2162)
; OTHER INFORMATION:
US-10-314-076-1

Query Match      100.0%; Score 1119; DB 16; Length 2198;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  ATGGAGCACAGCAGCCACCTCGCAGCCACAGCTCGCTGTGTTGGTGGTCCCGGC 60
DB      1  ATGGAGCACAGCAGCCACCTCGCAGCCACAGCTCGCTGTGTTGGTGGTCCCGGC 1106

QY      61  TCGGCTCGGGTTCGGTTCCTGTCGCTGCTACTACAGCTCTTGTGTGCTCGGT 120
DB      1107 TCGGCTCGGGTTCGGTTCCTGTCGCTGCTACTACAGCTCTTGTGTGCTCGGT 1166

QY      121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAGAGAGAAG 180
DB      1167 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAGAGAGAAG 1226

QY      181 TCCTCTCAACTATCTCTTGACATCTCGTGTGCTGGACATCTTGTCTCTTTTCATA 240
DB      1227 TCCTCTCAACTATCTCTTGACATCTCGTGTGCTGGACATCTTGTCTCTTTTCATA 1286

QY      241 GTGTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCCC 300
DB      1287 GTGTTGTGGACTTCTGTGTGAAGATTTCATCTTGAACATGCAGATGCCTCAGGTCCCC 1346

QY      301 GACAAGATCATGAAGTGTGGAATTCTCATCTCCATCCACACTCCATATGAATTACTGTA 360
DB      1347 GACAAGATCATGAAGTGTGGAATTCTCATCTCCATCCACACTCCATATGAATTACTGTA 1406

QY      361 CGGTTAAACATTGACAGTATATCGCTGTCTGCGACCGCTCAAGTACACACGCTCTCA 420
DB      1407 CGGTTAAACATTGACAGTATATCGCTGTCTGCGACCGCTCAAGTACACACGCTCTCA 1466

QY      421 TACCCAGCCGCGACCCGGAAGTCATTGTGAAGTGTATACATCACCTGCTTCTGACCAAGC 480
DB      1467 TACCCAGCCGCGACCCGGAAGTCATTGTGAAGTGTATACATCACCTGCTTCTGACCAAGC 1526

QY      481 ATCCCTATTAATGTGTGGCCCAACATCTGGAATGAAGACTACATCAGCACTCTGTGCAT 540
DB      1527 ATCCCTATTAATGTGTGGCCCAACATCTGGAATGAAGACTACATCAGCACTCTGTGCAT 1586

QY      541 CAGGCTCTCATCTGATCCACTGTTTACCGGTACTCGGTGCTGCTGCTTCCATCTTCTTC 600
DB      1587 CAGGCTCTCATCTGATCCACTGTTTACCGGTACTCGGTGCTGCTGCTTCCATCTTCTTC 1646

QY      601 ATCTTGAACATCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTTCGTCTCCGTGGC 660
DB      121 TTACAGCAAAATATCTTGTGACAGTGTATCATCTCTCCAGCTGGTGGGAGAGAGAAG 180
```

```
DB      1647 ATCTTGAACATCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTTCGTCTCCGTGGC 1706
QY      661 TACTCCACGGGAGAGACCACCGCCATCTTGTTCACCATTTACCTTCATCTTTGGCCACACTT 720
DB      1707 TACTCCACGGGAGAGACCACCGCCATCTTGTTCACCATTTACCTTCATCTTTGGCCACACTT 1766
QY      721 TGGGCCCCCGGCATCATCATGATTTCTTTACACACCTCTATGGGGGCCCATCCAGAACCGC 780
DB      1767 TGGGCCCCCGGCATCATCATGATTTCTTTACACACCTCTATGGGGGCCCATCCAGAACCGC 1826
QY      781 TGGCTGGTACACATCATGTCCGACATTTGCCAACATGTAGCCCTTTCTGAAACACAGCCATC 840
DB      1827 TGGCTGGTACACATCATGTCCGACATTTGCCAACATGTAGCCCTTTCTGAAACACAGCCATC 1886
QY      841 AACTTCTTCTTACTGCTTCAATGAGGAGGTTCCGACATGAGGAGCGCGCAGCGTC 900
DB      1887 AACTTCTTCTTACTGCTTCAATGAGGAGGTTCCGACATGAGGAGCGCGCAGCGTC 1946
QY      901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTTCTACCAATCAATAACTTTTCC 960
DB      1947 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTTCTACCAATCAATAACTTTTCC 2006
QY      961 ATACAAGTAGCCCTCGATCTCGCCGCAAACTCACACTGCATCAAGATGTGTGTAC 1020
DB      2007 ATACAAGTAGCCCTCGATCTCGCCGCAAACTCACACTGCATCAAGATGTGTGTAC 2066
QY      1021 CAGTATGACAAAAATGGAAAACTTATAAAAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 1080
DB      2067 CAGTATGACAAAAATGGAAAACTTATAAAAAAGTGTGTAATGACAGCAAAAGCTCTTACCAG 2126
QY      1081 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGCA 1119
DB      2127 TTTGAAGATGCCATTGGAGCTTGTGTATCATCTCTGTGCA 2165

RESULT 7
US-10-314-076-3
; Sequence 3, Application US/10314076
; Publication No. US20030152977A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRMY34, AND VARIANTS A
; FILE REFERENCE: D0197NP
; CURRENT APPLICATION NUMBER: US/10/314,076
; PRIOR FILING DATE: 2002-12-06
; PRIOR APPLICATION NUMBER: U.S. 60/338,371
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
; OTHER INFORMATION:
US-10-314-076-3

Query Match      97.5%; Score 1091; DB 16; Length 1110;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 1110; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

QY      1  ATGGAGCACAGCAGCCACCTCGCAGCCACAGCTCGCTGTCTTGTGTGTTCCCGGC 60
DB      1  ATGGAGCACAGCAGCCACCTCGCAGCCACAGCTCGCTGTCTTGTGTGTTCCCGGC 60

QY      61 TCGGCTCGGGCTGGGGTTTGTGCTGCTGTACTACAGCTCTTGTGTGCTCGGT 120
DB      61 TCGGCTCGGGCTGGGGTTTGTGCTGCTGTACTACAGCTCTTGTGTGCTCGGT 120

QY      121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAGAGAGAAG 180
DB      121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGGAGAGAGAAG 180
```

Db 121 TTACCGACCAATATCTGACAGTGATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
Qy 181 TCCTCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGTCTCTCTTTTCATA 240  
Db 181 TCCTCTACAACTATCTCTTGGCACTCGCTGCTGCCGACATCTTGTCTCTCTTTTCATA 240  
Qy 241 GTGTTGTGGACTCTCTTGTGGAAATTTTCATCTTGAACATGAGTGCCTCAGGTCCCG 300  
Db 241 GTGTTGTGGACTCTCTTGTGGAAATTTTCATCTTGAACATGAGTGCCTCAGGTCCCG 300  
Qy 301 GACAAGATCATGAAGTCTGGAATTTCTATCCATCCACACCTCCATATGGATTACTGTA 360  
Db 301 GACAAGATCATGAAGTCTGGAATTTCTATCCATCCACACCTCCATATGGATTACTGTA 360  
Qy 361 CGGTTAAACATTTGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACACGGTCTCA 420  
Db 361 CGGTTAAACATTTGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACACGGTCTCA 420  
Qy 421 TACCGACCGCGCACCGGAAAGTTCATTTGAAGTGTTCATACACCTGCTTCCCTGACGAG 480  
Db 421 TACCGACCGCGCACCGGAAAGTTCATTTGAAGTGTTCATACACCTGCTTCCCTGACGAG 480  
Qy 481 ATCCCTATTACTGTGGCCCAACATCTGCACTGGAAGTTCATACAGCACTCTGTGCAT 540  
Db 481 ATCCCTATTACTGTGGCCCAACATCTGCACTGGAAGTTCATACAGCACTCTGTGCAT 540  
Qy 541 CAGCTCTCATCTGGAATCACTGCTTCAACGCTTACCTGCTGCTCCATCTTCTTC 600  
Db 541 CAGCTCTCATCTGGAATCACTGCTTCAACGCTTACCTGCTGCTCCATCTTCTTC 600  
Qy 601 ATCTTGAATCAATATTTGTACAAGCTCAGGAGGAAGCAATTTTGGTCCGCTGTC 660  
Db 601 ATCTTGAATCAATATTTGTACAAGCTCAGGAGGAAGCAATTTTGGTCCGCTGTC 660  
Qy 661 TACTCCACGGGAGACACCGCATCTTGTTCACCATTTACCTCCATCTTTGGCACTT 720  
Db 661 TACTCCACGGGAGACACCGCATCTTGTTCACCATTTACCTCCATCTTTGGCACTT 720  
Qy 721 TGGGCCCCCGCATCATGATTTCTTTACCACTCTATGGGGGCGCCATCCAGAACCG 780  
Db 721 TGGGCCCCCGCATCATGATTTCTTTACCACTCTATGGGGGCGCCATCCAGAACCG 780  
Qy 781 TGCTGGTACACATCATGTCGCAATTTGCCAANGCTAGCCCTTCTGAACACAGCCATC 840  
Db 781 TGCTGGTACACATCATGTCGCAATTTGCCAANGCTAGCCCTTCTGAACACAGCCATC 840  
Qy 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCAACCATGGCAGCGCGCATC 900  
Db 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGCAACCATGGCAGCGCGCATC 900  
Qy 901 AAGGCTTTCTTCAAGTGCCAGAACCACTGTACAGTTCTTACACCAATCATAACTTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGCCAGAACCACTGTACAGTTCTTACACCAATCATAACTTTTCC 960  
Qy 961 ATAAACAAGTAGCCCTTGGATCTGCGCGCAAACTCACTGCACTCAAGATGCTGGTGATC 1020  
Db 961 ATAAACAAGTAGCCCTTGGATCTGCGCGCAAACTCACTGCACTCAAGATGCTGGTGATC 1020  
Qy 1021 CAGTATGACAAAATGGAAAACTTAAAGTGTGAATGACGAAAGCTCTTACCAG 1080  
Db 1021 CAGTATGACAAAATGG-----AAAAAGTGTGAATGACGAAAGCTCTTACCAG 1071  
Qy 1081 TTTGAAGATGCCATTGGAGCTTGTGTCATCATCTCTGTA 1119  
Db 1072 TTTGAAGATGCCATTGGAGCTTGTGTCATCATCTCTGTA 1110

RESULT 8

US-10-219-834-6

; Sequence 6, Application US/10219834

; Publication No. US20030096751A1

; GENERAL INFORMATION:

; APPLICANT: Bristol-Myers Squibb Company  
; TITLE OF INVENTION: G-PROTEIN COUPLED RECEPTOR POLYNUCLEOTIDES AND METHODS OF USE TH  
; FILE REFERENCE: D0191 NP  
; CURRENT APPLICATION NUMBER: US/10/219,834  
; PRIOR FILING DATE: 2002-08-15  
; PRIOR APPLICATION NUMBER: US 60/313,658  
; PRIOR FILING DATE: 2001-08-20  
; PRIOR APPLICATION NUMBER: US 60/340,703  
; PRIOR FILING DATE: 2001-10-30  
; PRIOR APPLICATION NUMBER: US 60/318,675  
; PRIOR FILING DATE: 2001-09-12  
; PRIOR APPLICATION NUMBER: US 60/355,596  
; PRIOR FILING DATE: 2002-02-06  
; PRIOR APPLICATION NUMBER: US 60/333,417  
; PRIOR FILING DATE: 2001-11-26  
; PRIOR APPLICATION NUMBER: US 60/338,367  
; PRIOR FILING DATE: 2001-12-06  
; NUMBER OF SEQ ID NOS: 192  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 6  
; LENGTH: 2189  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-10-219-834-6

Query Match 97.5%; Score 1091; DB 14; Length 2189;  
Best Local Similarity 99.2%; Pred. No. 0;  
Matches 1110; Conservative 0; Mismatches 0; Indels 9; Gaps 1;

Qy 1 ATGGAGCACACGACGCGCCACCTCGCAGCCACAGCTCGCTCTCTTGGTGGTCCCCGGC 60  
Db 1047 ATGGAGCACACGACGCGCCACCTCGCAGCCACAGCTCGCTCTCTTGGTGGTCCCCGGC 1106  
Qy 61 TCGCGCTGCGGCTTGGGTTTCGTGCCGCTGCTACTACAGCTCTTGTCTGTGCTCGCTCGGT 120  
Db 1107 TCGGCTGCGGCTTGGGTTTCGTGCCGCTGCTACTACAGCTCTTGTCTGTGCTCGGT 1166  
Qy 121 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 180  
Db 1167 TTACAGCAAAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAG 1226  
Qy 181 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGGGACATCTTGGTCTCTTTTCATA 240  
Db 1227 TCCTCTTACAACTATCTCTTGGCACTCGCTGCTGGGACATCTTGGTCTCTTTTCATA 1286  
Qy 241 GTGTTGTGGACTTCTGTTGGAAGATTTTCATCTTGAACATGCAATGCTCAGGTCCCC 300  
Db 1287 GTGTTGTGGACTTCTGTTGGAAGATTTTCATCTTGAACATGCAATGCTCAGGTCCCC 1346  
Qy 301 GACAAGATCATAGAAGTCTGGAAATTTCTCATCCATCCACCTCCATATGGATTACTGTA 360  
Db 1347 GACAAGATCATAGAAGTCTGGAAATTTCTCATCCATCCACCTCCATATGGATTACTGTA 1406  
Qy 361 CGGTTAAACATTTGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
Db 1407 CGGTTAAACATTTGACAGGTATATCGCTGCTGCCACCGCTCAAGTACCACAGGTCTCA 1466  
Qy 421 TACCGACCGCGCACCGGAAAGTTCATTTGAAGTGTTCATACACCTGCTTCCCTGACGAG 480  
Db 1467 TACCGACCGCGCACCGGAAAGTTCATTTGAAGTGTTCATACACCTGCTTCCCTGACGAG 1526  
Qy 481 ATCCCTATTACTGTGGCCCAACATCTGGAAGTTCATACAGCACTCTGTGTCAT 540  
Db 1527 ATCCCTATTACTGTGGCCCAACATCTGGAAGTTCATACAGCACTCTGTGTCAT 1586  
Qy 541 CAGCTCTCATCTGGATCACAATGCTTACCGTCTACCTGCTGCTGCTGCTGCTCTTCTTC 600  
Db 1587 CAGCTCTCATCTGGATCACAATGCTTACCGTCTACCTGCTGCTGCTGCTGCTCTTCTTC 1646  
Qy 601 ATCTTGAAGTCAATCATTTGTGTACAGCTCAGGAGGAGCAATTTTCTGCTCCGTGGC 660  
Db 1647 ATCTTGAAGTCAATCATTTGTGTACAGCTCAGGAGGAGCAATTTTCTGCTCCGTGGC 1706



PRIOR APPLICATION NUMBER: PCT/JP01/04643  
PRIOR FILING DATE: 2001-06-01  
PRIOR APPLICATION NUMBER: JP 2000-170446  
PRIOR FILING DATE: 2000-06-02  
PRIOR APPLICATION NUMBER: JP 2000-194926  
PRIOR FILING DATE: 2000-06-23  
NUMBER OF SEQ ID NOS: 8  
SEQ ID NO 3  
LENGTH: 1062  
TYPE: DNA  
ORGANISM: Human  
US-10-296-294A-3

Query Match 93.8%; Score 1050; DB 17; Length 1062;  
Best Local Similarity 100.0%; Pred. No. 7.3e-310;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGAGCACACGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGTCCCGCGC 60  
DB 1 ATGGAGCACACGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGTCCCGCGC 60  
QY 61 TCGGCTCGCGCTTGGGTTTGGCCGGTGTCTACTACAGCTCTTGTGTGCTCGGT 120  
DB 61 TCGGCTCGCGCTTGGGTTTGGCCGGTGTCTACTACAGCTCTTGTGTGCTCGGT 120  
QY 121 TTACCAAGCAAAATATTTGACAGTGATCATCTCTCCAGCTGGTGCAAGAGACAGAAG 180  
DB 121 TTACCAAGCAAAATATTTGACAGTGATCATCTCTCCAGCTGGTGCAAGAGACAGAAG 180  
QY 181 TCCTCTACAACTATCTTTGGGACTCGCTGTGCGGACATCTTGTGCTCTTTTTCATA 240  
DB 181 TCCTCTACAACTATCTTTGGGACTCGCTGTGCGGACATCTTGTGCTCTTTTTCATA 240  
QY 241 GTGTTTGTGACTTCTGTTGGAAGATTTTCATCTGCAACATGAGATGCTCAGGTCCCC 300  
DB 241 GTGTTTGTGACTTCTGTTGGAAGATTTTCATCTGCAACATGAGATGCTCAGGTCCCC 300  
QY 301 GACAAGATCATAGAGTGTGGAATTTCTCATCCATCCACACCTCCATATGGAATTA 360  
DB 301 GACAAGATCATAGAGTGTGGAATTTCTCATCCATCCACACCTCCATATGGAATTA 360  
QY 361 CGTTAAACATTTGACAGGTATATCGCTGTGCGGCTGCAAGTACACAGGTCTCA 420  
DB 361 CGTTAAACATTTGACAGGTATATCGCTGTGCGGCTGCAAGTACACAGGTCTCA 420  
QY 421 TACCCAGCCGACCCGGAAGTCAATGTAAGTGTTCATACACCTGCTTCTGACACG 480  
DB 421 TACCCAGCCGACCCGGAAGTCAATGTAAGTGTTCATACACCTGCTTCTGACACG 480  
QY 481 ATCCCTTATTTACTGTGGGCCCAACATCTGGAAGTCAATGTAAGTGTTCATACAC 540  
DB 481 ATCCCTTATTTACTGTGGGCCCAACATCTGGAAGTCAATGTAAGTGTTCATACAC 540  
QY 541 CAGTCTCTATCTGGAATCACTGTCTTCAAGTCTACCTGCTGCTGCTGCTGCTTTC 600  
DB 541 CAGTCTCTATCTGGAATCACTGTCTTCAAGTCTACCTGCTGCTGCTGCTGCTTTC 600  
QY 601 ATCTTGAATCAATCAATTTGTACAGCTCAGGAGAGCAATTTTCTGCTCCGTG 660  
DB 601 ATCTTGAATCAATCAATTTGTACAGCTCAGGAGAGCAATTTTCTGCTCCGTG 660  
QY 661 TACTTCAACGCGGAGACACCGCATCTTGTTCACCAATTTACCTTCCATCTTTGCA 720  
DB 661 TACTTCAACGCGGAGACACCGCATCTTGTTCACCAATTTACCTTCCATCTTTGCA 720  
QY 721 TGGGCCCCCGCATCATATGATTTTATACCACTCTATGCGGCGCCCATCCAGAAC 780  
DB 721 TGGGCCCCCGCATCATATGATTTTATACCACTCTATGCGGCGCCCATCCAGAAC 780  
QY 781 TGGCTGGTACATCATATGATTTTATACCACTCTATGCGGCGCCCATCCAGAAC 840  
DB 781 TGGCTGGTACATCATATGATTTTATACCACTCTATGCGGCGCCCATCCAGAAC 840

QY 841 AACTTCTTCTCTACTGTCTTCAATCAGCAAGCGGTTCCGCAACCATGCGAGCGCCAGCTC 900  
DB 841 AACTTCTTCTCTACTGTCTTCAATCAGCAAGCGGTTCCGCAACCATGCGAGCGCCAGCTC 900  
QY 901 AAGGCTTCTTCAAGTGGCCAGAGCAACCTGTACAGTCTTACACCAATCATTAACCTTTCC 960  
DB 901 AAGGCTTCTTCAAGTGGCCAGAGCAACCTGTACAGTCTTACACCAATCATTAACCTTTCC 960  
QY 961 ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1020  
DB 961 ATAACAAGTAGCCCTTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC 1020  
QY 1021 CAGTATGACAAAATGGAACCTTATAAA 1050  
DB 1021 CAGTATGACAAAATGGAACCTTATAAA 1050

RESULT 11

US-10-094-417-1  
; Sequence 1, Application US/10094417  
; Publication No. US20030045685A1  
; GENERAL INFORMATION:  
; APPLICANT: Tian, Hui  
; APPLICANT: Zhao, Jiagang  
; APPLICANT: Chen, Jin-Long  
; APPLICANT: Cutler, Gene  
; APPLICANT: Tularik Inc.  
; TITLE OF INVENTION: No. US20030045685A1el Receptors  
; FILE REFERENCE: 018781-008110US  
; CURRENT APPLICATION NUMBER: US/10/094,417  
; PRIOR FILING DATE: 2002-06-10  
; PRIOR APPLICATION NUMBER: US 09/802,803  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: US 60/276,649  
; PRIOR FILING DATE: 2001-03-16  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 1  
; LENGTH: 1202  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (68)..(1129)  
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) TGR20  
US-10-094-417-1

Query Match 93.8%; Score 1050; DB 14; Length 1202;  
Best Local Similarity 100.0%; Pred. No. 7.8e-310;  
Matches 1050; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGGAGCACACGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGTCCCGCGC 60  
DB 68 ATGGAGCACACGCGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGTCCCGCGC 127  
QY 61 TCGGCTCGCGCTTGGGTTTGGCCGGTGTCTACTACAGCTCTTGTGTGCTCGGT 120  
DB 128 TCGGCTCGCGCTTGGGTTTGGCCGGTGTCTACTACAGCTCTTGTGTGCTCGGT 187  
QY 121 TTACCAAGCAAAATATTTGACAGTGATCATCTCTCCAGCTGGTGCAAGAGACAGAAG 180  
DB 188 TTACCAAGCAAAATATTTGACAGTGATCATCTCTCCAGCTGGTGCAAGAGACAGAAG 247  
QY 181 TCCTCTACAACTATCTTTGGGACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240  
DB 248 TCCTCTACAACTATCTTTGGGACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 307  
QY 241 GTGTTTGTGAGACTTCTGTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCC 300  
DB 308 GTGTTTGTGAGACTTCTGTTGGAAGATTTTCATCTTGAACATGAGATGCTCAGGTCC 367







QY 601 ATCTTGAACATCAATCTGTTACAGCTCAGGAGGAGGAGCAATTTTGGTCTCCGTGGC 660  
Db 601 ATCTTGAACATCAATCTGTTACAGCTCAGGAGGAGGAGCAATTTTGGTCTCCGTGGC 660  
QY 661 TACTCCACGGGGAAGACACCGCATCTTGTTCACCAATACCTCCCATCTTTGGCCACATT 720  
Db 661 TACTCCACGGGGAAGACACCGCATCTTGTTCACCAATACCTCCCATCTTTGGCCACATT 720  
QY 721 TGGGCCCCCGGCATCATGATTTTACCACCTCTATGGGGCGCCCATCCAGAAACCGC 780  
Db 721 TGGGCCCCCGGCATCATGATTTTACCACCTCTATGGGGCGCCCATCCAGAAACCGC 780  
QY 781 TGGCTGGTACACATCATGTCGACATGCGCAACATGCTAGCCCTTCTGAACACAGCCATC 840  
Db 781 TGGCTGGTACACATCATGTCGACATGCGCAACATGCTAGCCCTTCTGAACACAGCCATC 840  
QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCGACACCATGCGCGCCGACGCTC 900  
Db 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCGACACCATGCGCGCCGACGCTC 900  
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960  
Db 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC 960  
QY 961 ATAAAGTAGCCCTTGGATCTCGCGGCAAACTCACTGCATCAAGATGCTGGGTGAC 1020  
Db 961 ATAAAGTAGCCCTTGGATCTCGCGGCAAACTCACTGCATCAAGATGCTGGGTGAC 1020  
QY 1021 CAGTATGACAAAATGGAACCTTATAAA 1050  
Db 1021 CAGTATGACAAAATGGAACCTTATAAA 1050

RESULT 13  
US-09-995-225-15  
; Sequence 15, Application US/09995225  
; Publication No. US20030139588A9  
; GENERAL INFORMATION:  
; APPLICANT: Chen, Ruoping  
; APPLICANT: Chu, Zhi Liang  
; APPLICANT: Dang, Huang T.  
; APPLICANT: Lowitz, Kevin P.  
; APPLICANT: Prigbe, Cameron  
; TITLE OF INVENTION: Endogenous And No. US20030139588A9-Endogenous Versions of Human G  
; TITLE OF INVENTION: Receptors  
; FILE REFERENCE: AREN-0308  
; CURRENT APPLICATION NUMBER: US/09/995,225  
; CURRENT FILING DATE: 2001-11-26  
; PRIOR APPLICATION NUMBER: 09/170,496  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: PCT/US99/23938  
; PRIOR FILING DATE: 1998-10-13  
; PRIOR APPLICATION NUMBER: 60/253,404  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/255,366  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 60/270,286  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/282,365  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/270,266  
; PRIOR FILING DATE: 2001-02-20  
; PRIOR APPLICATION NUMBER: 60/282,032  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/282,358  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/282,356  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: 60/290,917  
; PRIOR FILING DATE: 2001-05-14  
; PRIOR APPLICATION NUMBER: 60/309,208  
; PRIOR FILING DATE: 2001-07-31  
; NUMBER OF SEQ ID NOS: 67

; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 15  
; LENGTH: 1062  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: No. US20030139588A9e1 Sequence  
US-09-995-225-15

Query Match 93.7%; Score 1048.4; DB 10; Length 1062;  
Best Local Similarity 99.9%; Pred. No. 2.3e-309;  
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGAGCACACGACGCCACCTCGCAGCCAAACAGCTCGCTGCTTTGGTGGTCCCGCGC 60  
Db 1 ATGAGAGCACACGACGCCACCTCGCAGCCAAACAGCTCGCTGCTTTGGTGGTCCCGCGC 60  
QY 61 TCGGCTGGGGCTTGGGTTTCGTGCCGTGGTCTACTACAGCTCTTGTGTGGCTCGGT 120  
Db 61 TCGGCTGGGGCTTGGGTTTCGTGCCGTGGTCTACTACAGCTCTTGTGTGGCTCGGT 120  
QY 121 TTACCAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGAGACAGAAG 180  
Db 121 TTACCAGCAAAATATCTTGACAGTGTATCTCTCCAGCTGGTGGCAAGAGACAGAAG 180  
QY 181 TCCTCTTACAAATATCTTTGGCACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240  
Db 181 TCCTCTTACAAATATCTTTGGCACTCGCTGTGCGGACATCTTGGTCTCTTTTTCATA 240  
QY 241 GTGTTTGGGACTTCTGTGTGAAGATTTTCATCTTGAACATGACAGTGGCTCAGGTCCCC 300  
Db 241 GTGTTTGGGACTTCTGTGTGAAGATTTTCATCTTGAACATGACAGTGGCTCAGGTCCCC 300  
QY 301 GACAAAGATCATAGAAAGTGTGGAAATTTCTCATCATCCACCTCCCATATGGATTAAGTGA 360  
Db 301 GACAAAGATCATAGAAAGTGTGGAAATTTCTCATCATCCACCTCCCATATGGATTAAGTGA 360  
QY 361 CCGTTAAACCAATGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
Db 361 CCGTTAAACCAATGACAGGTATATCGCTGTCTGCCACCGCTCAAGTACCACAGGTCTCA 420  
QY 421 TACCAGCGCGCACCGGAAAGTCAATGTAAGTGTATACATCACCTGCTCTCTGACACG 480  
Db 421 TACCAGCGCGCACCGGAAAGTCAATGTAAGTGTATACATCACCTGCTCTCTGACACG 480  
QY 481 ATCCCTTATTACTGTGGCCCAACATCTGAGCTGAAGACTACATCAGACACCTCTGTGCAT 540  
Db 481 ATCCCTTATTACTGTGGCCCAACATCTGAGCTGAAGACTACATCAGACACCTCTGTGCAT 540  
QY 541 CAGCTCTCATCTGGATCCATCTGCTTCCACCGTCTACCTGGTGGCTCCCATCTTTCTTC 600  
Db 541 CAGCTCTCATCTGGATCCATCTGCTTCCACCGTCTACCTGGTGGCTCCCATCTTTCTTC 600  
QY 601 ATCTTGAACATCAATCTGTTGTAAGCTCAGAGGAGAGCAATTTTGGTCCGTGGC 660  
Db 601 ATCTTGAACATCAATCTGTTGTAAGCTCAGAGGAGAGCAATTTTGGTCCGTGGC 660  
QY 661 TACTCCAGGGGAAGACACCGCATCTTGTTCACCAATTTACCTCCCATCTTTGGCCACATT 720  
Db 661 TACTCCAGGGGAAGACACCGCATCTTGTTCACCAATTTACCTCCCATCTTTGGCCACATT 720  
QY 721 TGGGCCCCCGGCATCATGATTTTACCACCTCTATGGGGCGCCCATCCAGAAACCGC 780  
Db 721 TGGGCCCCCGGCATCATGATTTTACCACCTCTATGGGGCGCCCATCCAGAAACCGC 780  
QY 781 TGGCTGGTACACATCATGTCGACATGCGCAACATGCTAGCCCTTCTGAACACAGCCATC 840  
Db 781 TGGCTGGTACACATCATGTCGACATGCGCAACATGCTAGCCCTTCTGAACACAGCCATC 840  
QY 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCGACACCATGCGCGCCGACGCTC 900  
Db 841 AACTTCTTCTCTACTGCTTTCATCAGCAAGCGGTTCGACACCATGCGCGCCGACGCTC 900

QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCAATAACTTTTCC 960  
DB |||||||  
QY 901 AAGGCTTTCTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCAATAACTTTTCC 960  
DB |||||||  
QY 961 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACATGCGATCAAGATGCTGGTGATC 1020  
DB |||||||  
QY 961 ATAACAAGTAGCCCTGGATCTCGCGGCAAACTCACATGCGATCAAGATGCTGGTGATC 1020  
DB |||||||  
QY 1021 CAGTATGACAAAAATGGAACCTTATAAAA 1050  
DB |||||||  
QY 1021 CAGTATGACAAAAATGGAACCTTATAAAA 1050  
DB |||||||  
RESULT 14  
US-10-296-294A-4  
; Sequence 4, Application US/10296294A  
; Publication No. US20040029224A1  
; GENERAL INFORMATION:  
; APPLICANT: MATSUI, Hideki  
; APPLICANT: TERAU, Yasuko  
; APPLICANT: SHINTANI, Yasushi  
; FILE OF INVENTION: No. US20040029224A1el G Protein-Coupled Receptor and its DNA  
; FILE REFERENCE: 2734 USOP  
; CURRENT APPLICATION NUMBER: US/10/296,294A  
; CURRENT FILING DATE: 2001-11-21  
; PRIOR APPLICATION NUMBER: PCT/JP01/04643  
; PRIOR FILING DATE: 2001-06-01  
; PRIOR APPLICATION NUMBER: JP 2000-170446  
; PRIOR FILING DATE: 2000-06-02  
; PRIOR APPLICATION NUMBER: JP 2000-194926  
; PRIOR FILING DATE: 2000-06-23  
; NUMBER OF SEQ ID NOS: 8  
; SEQ ID NO 4  
; LENGTH: 1062  
; TYPE: DNA  
; ORGANISM: Human  
US-10-296-294A-4

Query Match 93.7%; Score 1048.4; DB 17; Length 1062;  
Best Local Similarity 99.9%; Pred. No. 2.3e-309;  
Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAGGACACAGCAGCCACCTCGCAGCAACAGCTCGCTGTCTTGTGGTCCCGCGC 60  
DB |||||||  
QY 1 ATGAGGACACAGCAGCCACCTCGCAGCAACAGCTCGCTGTCTTGTGGTCCCGCGC 60  
DB |||||||  
QY 61 TCGGCTCGCGCTTGGGTTTGGTGGCCGCTGTACTACAGCTCTTGTGCTCGGT 120  
DB |||||||  
QY 121 TTACAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
DB |||||||  
QY 121 TTACAGCAATATCTTGACAGTATCATCTCTCCAGCTGGTGGCAAGACAGAAAG 180  
DB |||||||  
QY 181 TCCTCCTACAATCTCTTGGCACTCGCTGCTGCGGACATCTTGCTCTCTTTTTCATA 240  
DB |||||||  
QY 181 TCCTCCTACAATCTCTTGGCACTCGCTGCTGCGGACATCTTGCTCTCTTTTTCATA 240  
DB |||||||  
QY 241 GTGTTTGTGGACTTCTCTGTGGAAGATTTCTATCTTTGAACATGCGATGCTCGCTCCC 300  
DB |||||||  
QY 301 GACAAGATCATGAAGTGTGAATTTCTATCCATCCACCTCATATGATGATTAAGTGA 360  
DB |||||||  
QY 301 GACAAGATCATGAAGTGTGAATTTCTATCCATCCACCTCATATGATGATTAAGTGA 360  
DB |||||||  
QY 361 CGGTTAACATTGACAGGTATATCGCTGTGSCACCGCTCAAGTACCACAGCTCTCA 420  
DB |||||||  
QY 361 CGGTTAACATTGACAGGTATATCGCTGTGSCACCGCTCAAGTACCACAGCTCTCA 420  
DB |||||||  
QY 421 TACCAGCCCGCAGCCGGAAGTCAATGTGAAGTGTATACATCAGCTGCTTCTGACACAGC 480  
DB |||||||  
QY 421 TACCAGCCCGCAGCCGGAAGTCAATGTGAAGTGTATACATCAGCTGCTTCTGACACAGC 480  
DB |||||||

QY 481 ATCCCTATTACTGTTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
DB |||||||  
QY 481 ATCCCTATTACTGTTGGCCCAACATCTGGACTGAAGACTACATCAGCACCTCTGTGCAT 540  
DB |||||||  
QY 541 CACGCTCTCATCTGGATCCACTGCTTACCCGCTTACCTGTGGTGGCCCTGTCCATCTTTTC 600  
DB |||||||  
QY 541 CACGCTCTCATCTGGATCCACTGCTTACCCGCTTACCTGTGGTGGCCCTGTCCATCTTTTC 600  
DB |||||||  
QY 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTCTGCTCCGTCGC 660  
DB |||||||  
QY 601 ATCTTGAACTCAATCATTTGTGTACAAGCTCAGGAGGAAGAGCAATTTTCTGCTCCGTCGC 660  
DB |||||||  
QY 661 TACTCCAGGGGAGAGCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720  
DB |||||||  
QY 661 TACTCCAGGGGAGAGCAACCGCCATCTTGTTCACCATTTACCTCCATCTTTGCCACACTT 720  
DB |||||||  
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGGGCCCATCCAGAACCGC 780  
DB |||||||  
QY 721 TGGGCCCCCGCATCATCATGATTTCTTTACCACTCTATGGGGGGCCCATCCAGAACCGC 780  
DB |||||||  
QY 781 TGGTGTGTACATCATGTTCGACATTTGCCAATTCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
DB |||||||  
QY 781 TGGTGTGTACATCATGTTCGACATTTGCCAATTCGCAACATGTAGCCCTTCTGAAACACAGCCATC 840  
DB |||||||  
QY 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCGCGCCTC 900  
DB |||||||  
QY 841 AACTTCTTCTCTACTGCTTCATCAGCAAGCGGTTCCGACCATGGCAGCGCCGCGCCTC 900  
DB |||||||  
QY 901 AAGGCTTTCTTCAAGTGCAGAAACCTGTACAGTTCTACACCAATCAATAACTTTTTC 960  
DB |||||||  
QY 901 AAGGCTTTCTTCAAGTGCAGAAACCTGTACAGTTCTACACCAATCAATAACTTTTTC 960  
DB |||||||  
QY 961 ATAAACAAGTAGCCCTCGATCTCGCGGCAAACTCACATGCGATCAAGATGCTGGTGATC 1020  
DB |||||||  
QY 961 ATAAACAAGTAGCCCTCGATCTCGCGGCAAACTCACATGCGATCAAGATGCTGGTGATC 1020  
DB |||||||  
QY 1021 CAGTATGACAAAAATGGAACCTTATAAAA 1050  
DB |||||||  
QY 1021 CAGTATGACAAAAATGGAACCTTATAAAA 1050  
DB |||||||

## RESULT 15

US-09-813-432-21  
; Sequence 21, Application US/09813432  
; Publication No. US20030148485A1  
; GENERAL INFORMATION:  
; APPLICANT: Taupier Jr., Raymond J  
; APPLICANT: Majmuder, Kamud  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Mezes, Peter S  
; APPLICANT: Vernet, Corine A. M.  
; TITLE OF INVENTION: No. US20030148485A1el Polypeptides and Amino Acids Encoding Same  
; FILE REFERENCE: 15966-729  
; CURRENT APPLICATION NUMBER: US/09/813,432  
; CURRENT FILING DATE: 2001-03-20  
; PRIOR APPLICATION NUMBER: 60/190,835  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,768  
; PRIOR FILING DATE: 2000-03-20  
; PRIOR APPLICATION NUMBER: 60/190,972  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,199  
; PRIOR FILING DATE: 2000-03-22  
; PRIOR APPLICATION NUMBER: 60/191,947  
; PRIOR FILING DATE: 2000-03-24  
; PRIOR APPLICATION NUMBER: 60/192,665  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,657  
; PRIOR FILING DATE: 2000-03-28  
; PRIOR APPLICATION NUMBER: 60/192,984  
; PRIOR FILING DATE: 2000-03-28

		; PRIOR APPLICATION NUMBER: 60/192,664	
		; PRIOR FILING DATE: 2000-03-28	
		; PRIOR APPLICATION NUMBER: 60/192,836	
		; PRIOR FILING DATE: 2000-03-29	
		; PRIOR APPLICATION NUMBER: 60/193,843	
		; PRIOR FILING DATE: 2000-03-31	
		; NUMBER OF SEQ ID NOS: 78	
		; SOFTWARE: Patentin Ver. 2.1	
		; SEQ ID NO 21	
		; LENGTH: 1343	
		; TYPE: DNA	
		; ORGANISM: Homo sapiens	
		US-09-813-432-21	
		Query Match 93.7%; Score 1048.4; DB 10; Length 1343;	
		Best Local Similarity 99.9%; Pred. No. 2.6e-309;	
		Matches 1049; Conservative 0; Mismatches 1; Indels 0; Gaps 0;	
QY	1	ATGGAGCACAGCAGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGGTCCCGGC	60
DB	2	ATGGAGCACAGCAGCCACCTCGCAGCCAAACAGCTCGCTGTCTTGGTGGTCCCGGC	61
QY	61	TCGGCTCGGGCTTGGGTTTCGTGCCCGTGTCTACTACAGCTCTTGTGTGCTCGGT	120
DB	62	TCGGCTCGGGCTTGGGTTTCGTGCCCGTGTCTACTACAGCTCTTGTGTGCTCGGT	121
QY	121	TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGACAGAAAG	180
DB	122	TTACCAGCAATATCTTGACAGTGATCATCTCTCCAGCTGTGGCAAGACAGAAAG	181
QY	181	TCCTCTCAACATATCTTTGGCAGTCGCTGTGCGCAATCTTGGTCTCTTTTTCATA	240
DB	182	TCCTCTCAACATATCTTTGGCAGTCGCTGTGCGCAATCTTGGTCTCTTTTTCATA	241
QY	241	GTGTTGTGGATCTTCTGTGGAAGATTTCACTTGAACATGCAGATGCTCAGGTCCCC	300
DB	242	GTGTTGTGGATCTTCTGTGGAAGATTTCACTTGAACATGCAGATGCTCAGGTCCCC	301
QY	301	GACAGATCATAGAGTGTGGAATTTCTCATCATCCACCTCCATATGGATTACTGTA	360
DB	302	GACAGATCATAGAGTGTGGAATTTCTCATCATCCACCTCCATATGGATTACTGTA	361
QY	361	CGGTTAACCATTTGACAGGTATATCGCTGTGCGCACCCGCTCAAGTACACACGGTCTCA	420
DB	362	CGGTTAACCATTTGACAGGTATATCACTGTGCGCACCCGCTCAAGTACACACGGTCTCA	421
QY	421	TACCCAGCCGCGACCCGGAAGTCATTGTAAGTTTATCATCCTGCTTCTGACCAGC	480
DB	422	TACCCAGCCGCGACCCGGAAGTCATTGTAAGTTTATCATCCTGCTTCTGACCAGC	481
QY	481	ATCCCCCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGCAT	540
DB	482	ATCCCCCTATTACTGTGGGCCCAACATCTGGACTGGAAGACTACATCAGCACCTCTGTGCAT	541
QY	541	CAGTCTCTCATCTGGATCACTGCTTTACCGTCTACCTGGTGGCCCTGCTCCATCTTTTC	600
DB	542	CAGTCTCTCATCTGGATCACTGCTTTACCGTCTACCTGGTGGCCCTGCTCCATCTTTTC	601
QY	601	ATCTTGAATCATCATCTGTGTACAGCTCAGGAGGAAGCAATTTTCGTCTCCGTGGC	660
DB	602	ATCTTGAATCATCATCTGTGTACAGCTCAGGAGGAAGCAATTTTCGTCTCCGTGGC	661
QY	661	TACTCCAGGGGAAGACACCGGCATCTTGTTCACCATTTACCTTCATCTTTGCCACACTT	720
DB	662	TACTCCAGGGGAAGACACCGGCATCTTGTTCACCATTTACCTTCATCTTTGCCACACTT	721
QY	721	TGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC	780
DB	722	TGGGCCCCCGCATCATCATGATTTCTTTTACCACTCTATGGGGCGCCCATCCAGAACCGC	781
QY	781	TGGCTGGTACACATCATGTCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC	840
DB	782	TGGCTGGTACACATCATGTCGACATTTGCCAATGCTAGCCCTTCTGAACACAGCCATC	841

Search completed: February 12, 2005, 13:39:46  
Job time : 662 secs

QY	841	AACTTCTTCTCTACTGTCTTCAACAGCAAGCGTTCCGCAACCATGGCAGCCGCGCTC	900
DB	842	AACTTCTTCTCTACTGTCTTCAACAGCAAGCGTTCCGCAACCATGGCAGCCGCGCTC	901
QY	901	AAGGCTTTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC	960
DB	902	AAGGCTTTTCAAGTGCAGAGCAACCTGTACAGTTCTACACCAATCATAACTTTTCC	961
QY	961	ATAACAAGTAGCCCTCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC	1020
DB	962	ATAACAAGTAGCCCTCTGGATCTCGCGGCAAACTCACACTGCATCAAGATGCTGGTGTAC	1021
QY	1021	CAGTATGACAAAATGGAACCTATAAAA	1050
DB	1022	CAGTATGACAAAATGGAACCTATAAAA	1051

**THIS PAGE BLANK (USPTO)**